INDEX NUMBERS ELUCIDATED

Willford I. King

Longmans' Economics Series

Digitized by the Internet Archive
in 2023 with funding from
Kahle/Austin Foundation

https://archive.org/details/indexnumberseluc0000will

INDEX NUMBERS ELUCIDATED

Longmans' Economics Series

GENERAL EDITOR

ERNEST L. BOGART

PROFESSOR OF ECONOMICS IN THE
UNIVERSITY OF ILLINOIS

Economic History of American Agriculture. *By* ERNEST L. BOGART.

Readings in the Economic History of the United States. *By* ERNEST L. BOGART *and* CHARLES M. THOMPSON, *University of Illinois.*

Modern Industry. *By* ERNEST L. BOGART *and* CHARLES E. LANDON, *Duke University.*

Money. *By* FREDERICK A. BRADFORD, *Lehigh University.*

Taxation in the Modern State. *By* ALZADA COMSTOCK, *Mount Holyoke College.*

Interpretive Accounting. *By* F. E. FOLTS *and* A. B. STILLMAN, *University of Oregon.*

American Monetary and Banking Policies. *By* GEORGE W. DOWRIE, *Leland Stanford University.*

Index Numbers Elucidated. *By* WILLFORD I. KING, *New York University.*

IN PREPARATION

International Trade. *By* H. T. COLLINGS, *University of Pennsylvania.*

Business Ethics. *By* A. D. KAPLAN, *University of Denver.*

The Modern Railroad. *By* JULIUS A. PARMELEE of Washington, D. C.

Money and Banking. *By* W. E. SPAHR, *New York University.*

Industrial History of the United States. *By* ERNEST L. BOGART.

Longmans' Economics Series

INDEX NUMBERS ELUCIDATED

BY

WILLFORD ISBELL KING, PH.D

PROFESSOR OF ECONOMICS, NEW YORK UNIVERSITY

UNIVERSITY OF CALIFORNIA LIBRARY

LONGMANS, GREEN AND CO.

NEW YORK · LONDON · TORONTO

1930

KING

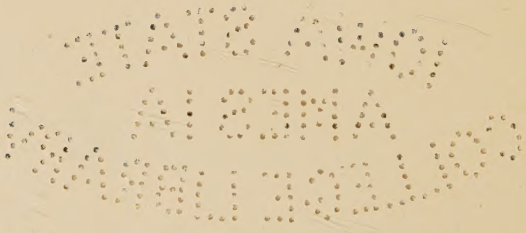
INDEX NUMBERS ELUCIDATED

COPYRIGHT · 1930

BY LONGMANS, GREEN AND CO.

ALL RIGHTS RESERVED, INCLUDING THE RIGHT TO REPRODUCE THIS BOOK, OR ANY PORTION THEREOF, IN ANY FORM

FIRST EDITION



PRINTED IN THE UNITED STATES OF AMERICA

HB225
K589i
c.2

TO

MY NEVER-FAILING SOURCE OF INSPIRATION

JANE PATTERSON KING

231803

PREFACE

WHY another book on index numbers? Is there anything worth while saying concerning this topic that has not already been said? The answer to these questions is that, while existing treatises in this field are excellent from the descriptive standpoint, no one of them which has come to the attention of the present author either brings out adequately the real significance of index numbers or presents the correct criterion for judging of the merits of these devices and illustrates its use. Furthermore, most of the treatises dealing with the nature, construction, and interpretation of index numbers either cover but one or two phases of the subject, or are so technical in style that they are difficult reading for all except mathematical experts.

At present, however, practically every educated man or woman finds it imperative to employ index numbers, and it seems highly desirable that the users should know something of the real nature of these devices. Furthermore, there are today in our colleges and universities many thousands of students of economics and statistics who are anxious to ascertain the correct answer to the following questions:

1. What do the various types of index numbers mean? How are they to be interpreted? How do they differ from one another? When should each be used?
2. What are the actual mechanics of constructing an index number of a given type?

The average student doubtless prefers to have the answer

to the last question set forth in plain arithmetic. He wants to know the convenient way of arranging the figures when an index number is to be computed. Existing discussions of the subject do not illustrate the processes in a way simple enough to meet his needs, for most of the writers either assume that the process is too simple to need illustrating, or have felt compelled to fill their books with algebraic formulae which, while simple enough for mathematicians, appall the ordinary undergraduate.

It seems, then, that there is room for a book designed to serve the needs of that great majority of persons, in college and outside, who must use index numbers—the majority who do not have a mathematical turn of mind and who detest algebraic formulae. This work is intended primarily to meet the demands of this great class.

It would be a decided mistake, however, to infer that the present book goes no further than to re-state facts, and to re-describe processes already well known and generally accepted—for nothing could be further from the truth. Such a text would necessarily be diminutive in size, owing to the fact that there are so few points upon which all authorities agree. Furthermore, most of the treatises on index numbers err lamentably in one or more of the following respects:

1. Too sharp a dividing line is drawn between index numbers derived by averaging relatives and index numbers based on aggregates, the writers failing to give due recognition to the fact that as a rule the differences between the two methods of obtaining an index number can easily be resolved into nothing more than a change in weights.
2. The assumption is made that all index numbers are designed primarily for a single purpose, that purpose being to answer the question: *What changes have occurred in the general purchasing power of money?*

3. It has been assumed that the failure of an index series to conform to the central tendency of a large group of index numbers is an evidence of "bias," in other words, that index numbers computed from a given set of sample data by one formula tend to climb too rapidly, while, if computed by another formula, there is a tendency to fall too low. The truth seems to be that conformity to a central tendency is far from being adequate proof that an index number is accurate. Furthermore, in many cases, such "bias" as exists is not at all a characteristic of the formula, but arises instead either from the peculiarity of the data dealt with, or from poor sampling.
4. The hypothesis is set forth that there exists an ideal formula for computing an index number, and that the index number calculated by use of this formula must be the best for all purposes.
5. It is taken for granted that the mathematical formula correctly applicable to the *totality* or *universe of items* must likewise necessarily be the best formula for use in the construction of an index number intended to represent that totality.
6. It is assumed that the merit of an index number can be judged by studying its formula to see how well that formula meets the following requirements:

    (a) The circular test
    (b) The factor-reversal test
    (c) The commodity-reversal test
    (d) The time-reversal test

7. The important fact that an index number is nothing but a particular type of sample, and that, therefore, the problem of index numbers is merely one phase of the general problem of sampling, is often either

entirely overlooked or treated as a matter of no particular significance.

8. Because of the oversight just mentioned, few, if any, students of the subject seem to have realized that the only formula test which has any merit is: *Does this formula, when applied to a sample, give results, the fluctuations in which closely resemble those occurring in the totality of items?*

In view of the number and importance of the errors just enumerated, it appears that there is ample reason for explaining in the simplest language possible the nature of index numbers, how they are constructed, and how they must be interpreted. The author would be the last to contend that all previous writers have been deceived concerning all of these eight points, for the fact is that, in respect to points 1, 2, and 4, only a minority of writers on index numbers seem to be in error, and, presumably, some writer at some time and in some place, has stated or at least hinted at the truth in each of the other five points. It appears, however, that none of these investigators has ever published a study of index numbers which has correctly set forth with properly graduated emphasis the facts concerning all eight of these points; indeed, the present writer believes that no existing treatment of the subject gives both an approximately true analysis of the fundamentals of index numbers, and a correct statement of the relationship of these fundamentals to each other. It is in an effort to supply a proper orientation of the facts that this book has been written.

The general thesis here set forth differs so radically from those appearing in most standard texts that one is almost justified in saying that a new theory of index numbers is here presented. The question naturally arises as to whether this revised doctrine is an improvement over the ones which it seeks to displace. The author has attempted to analyze

each step so clearly that the reader will be compelled to take nothing on faith. But this is not all. He has submitted statistical evidence to cover as many as possible of the important questions of fact. The reader must be the judge as to whether or not the case, as presented in the following pages, is well established.

It should be clearly understood that this book is devoted to a study of principles and not of concrete facts, and that, therefore, any description of existing index numbers is purely incidental. Those interested in such descriptions are referred to Bulletin 284 of the United States Bureau of Labor Statistics, a bulletin in which Professor Wesley C. Mitchell sets forth the chief characteristics of the better-known published index numbers of prices, or to Professor Irving Fisher's forthcoming *Encyclopedia of Index Numbers*—a work which will contain brief descriptions of a large number of index numbers.

The writer wishes to acknowledge his indebtedness to the National Bureau of Economic Research, which furnished the statistical clerical help used in the preliminary stages of this study, to Dr. Ernest L. Bogart, Mr. Walter R. Peabody, and Dr. Maurice Leven for numerous very useful suggestions, to Dr. Frederick C. Mills for his many helpful criticisms, to Mr. Charles K. Payne and Miss Lucile Bagwell for aid in preparing the bibliographical notes, to Misses Harriet Dobson, Bertha Jacobson, and Betty Dublin, who deserve much credit for their efficient work in making the computations for the tables, and to Mr. Benjamin S. Levin and Miss Rose Dessel who have added in the preparation of the manuscript for the printer.

Willford I. King

CONTENTS

CHAPTER I

PAGE

INTRODUCTION 1

CHAPTER II

RELATIVE COMPARISONS USING COMPLETE DATA..... 5

More interest in relative than in absolute comparisons... 5

The conversion of absolute numbers to relatives..... 5

The advantage of relatives over absolute numbers..... 7

The formulation of mathematical answers to questions relating to degree of change..... 7

The base defined..... 8

The usefulness of relatives not confined to time series.... 10

The choice of a base..... 11

Ratio charts 18

Relative comparisons representing a group of series.... 18

Different relatives required to answer different questions.. 19

A change in base equivalent to a change in weights..... 23

With a common base the weighted aggregate may equal the simple arithmetic mean of relatives..... 25

The question asked governs the formula required for obtaining the answer..... 25

Relative aggregate values at various dates of a fixed number of shares of stock..... 31

In comparisons based upon aggregates of original items the larger items dominate the results..... 31

PAGE

Relationship between the simple aggregate of original data and the weighted arithmetic mean of relatives.... 32
Possible effect of using one base for prices and another for weights 35
In computing price-relatives it is essential to compare the same quantities of identical commodities..... 36
Difficulty of comparing the prices of different budgets... 37
Commodity substitution in price comparisons..... 38
Composite relatives 39
With complete data exact answers are possible..... 40
Summary 41

CHAPTER III

The Nature of Index Numbers..... 45
Why the subject of index numbers is ill understood..... 45
Index numbers of special class of relatives..... 46
Index numbers defined..... 47
The problem of index numbers..... 49
Index numbers only useful to answer precise questions... 49
Different formulæ are required to answer different questions 51
There can be no "best" formula for an index number.... 53
Supposed tests of the quality of index numbers..... 55
Summary 56

CHAPTER IV

Sampling as Related to Index Numbers..... 59
Complete data versus sample data..... 59
The problem of sampling..... 59
Accuracy of sample items..... 60
Average prices 61
The number of items required for a satisfactory sample.. 62
Random versus proportional sampling..... 64

PAGE
The weighting of the sample items 67
The method of weighted aggregates 68
The weighted arithmetic average of relatives 69
Advantages inhering in the average of relatives 69
The introduction of new items into a sample composed of relatives 71
What weights are appropriate? 72
Continuous data may constitute misleading sample 73
Summary 75

CHAPTER V

INDEX NUMBERS DERIVED FROM LINK RELATIVES OR RATIOS 78
Two different kinds of chain comparisons 78
Use of link relatives forced by lack of continuous, comparable data 79
The problem raised by changing the composition of the sample 80
Linking the simple averages of relatives 81
Linking the weighted averages of relatives 81
Average of relatives may be so weighted as to give results identical with those obtained by the method of aggregates 82
Use of the method of aggregates with a changing sample. 86
Index numbers derived by splicing ratios obtained by method of aggregates 86
Characteristics of indices obtained by linking averages of relatives 89
1921–1922; equal fractional changes up and down 89
1922–1923; equal absolute changes up and down 90
1923–1924; inverse proportional changes up and down in commodities having identical values 91
1924–1925; high-priced commodities increasing and low-priced commodities diminishing in value 91

PAGE

1925–1926; high-priced commodities falling and low-priced commodities rising 92
1926–1927; return of the price of every commodity to its original level 94
The desirability of studying index numbers by experimenting with actual data 94
The stock-price records 95
The selection of three samples 96
Computation of indices from fixed-base and from chain relatives 97
Comparative trends of fixed-base and chained relatives of stock prices 98
Commodity price data used in this analysis of index numbers 100
Selection of samples from which to compute indices of commodity prices 101
Calculation of fixed-base and link relatives of commodity prices 101
Comparative trends of indices of commodity prices derived from fixed-base relatives and from link relatives .. 101
Summary 102

CHAPTER VI

TYPES AND AVERAGES AS RELATED TO INDEX NUMBERS. 106
Need for understanding types and averages 106
Types and averages used in relative comparison 106
Relative sizes of the arithmetic, geometric, and harmonic averages 110
Geometric average not affected by changing the base for the individual relatives 111
Explanation of the supposed "bias" of certain index numbers 113
Arithmetic, geometric, and harmonic averages of original data compared 114

PAGE
Converting items to relatives is equivalent to giving new weights to the items 115
The trends of five different kinds of index numbers computed from actual prices of stocks 116
Computation of aggregates and relatives of stock prices. 119
Diversity in the movements of different classes of stocks.. 119
Reasons for variance in behavior of stock-price index numbers computed by different formulæ 124
Trend of weighted average of relatives dependent upon weights used 128
Computation of indices of commodity prices 129
Behavior of indices derived by different formulæ 130
Comparison of weighted and unweighted arithmetic averages 132
Simple and weighted arithmetic averages of relatives of various classes of commodities 135
Differences in trends of weighted and unweighted arithmetic averages of relatives due to chance variations... 138
Summary 138

CHAPTER VII

PERCENTAGES OF ERROR FOUND IN CERTAIN PRICE INDICES 143
Experimentation necessarily limited in scope 143
Established facts concerning index numbers 144
Formula applicable to totality commonly assumed to be best for use with sample data 144
Stock-Prices per Share 145
Nature of stock-price data used in experiments 145
Average of totality not identical with average of universe 145
Deviations from the Aggregate Obtained by Adding the Prices Per Share of 302 Stocks 146
With complete data questions can be answered precisely.. 146
Reason for using samples instead of complete data 147
Testing index numbers for accuracy 147

PAGE
The relative merits of the indices tested..... 150
Increase of error with departure from base..... 150
Relation between the size of a sample and its percentage of error 151
Relation of errors to general "law" of error..... 157
Reducing the error by omitting the low-priced stocks.... 157
Commodity Price-Relatives 161
Deviations from the Simple Arithmetic Average of 336 Relatives of Commodity Prices..... 162
Errors of index numbers compared for stock prices and commodity prices 163
Comparative errors of different index numbers..... 164
Relation of size of error to time-distance from base.... 165
Relation of size of error to size of sample..... 166
Deviations from the Weighted Arithmetic Average of 336 Relatives of Commodity Prices..... 168
Methods used in constructing index numbers..... 169
Errors little affected by time-distance from base..... 171
Errors of different index series compared..... 171
Why weighting increases the error of the arithmetic average of a small sample..... 171
Comparative effects of individual weighting and group weighting 173
Arithmetic averages often vitiated by erratic movements of one or two price-relatives..... 176
Relation of size of error to size of sample..... 177
Comparative merits of random and proportional sampling 181
Improving the sample by eliminating erratic items..... 182
The best index number for answering question 3..... 183
Summary 184

CHAPTER VIII

MEASURING THE PURCHASING POWER OF MONEY..... 189
To measure the purchasing power of money is sometimes assumed to be the only important use of index numbers 189

PAGE
The chief uses of an index of the general purchasing power of money ... 190
Chief uses of money ... 191
Use for purchase of commodities is one most studied ... 191
The equation of exchange ... 192
Two sides of equation of exchange refer to different dates 193
The total value of goods sold ... 193
The physical volume of goods sold ... 194
The best quantity units for use in measuring the general purchasing power of money ... 195
Which base is best is a question impossible to decide ... 196
Measurement of physical volume of trade difficult because many important commodities are not standardized ... 199
General purchasing power of money must perforce be estimated from sample data ... 200
Sample should be widely inclusive ... 200
Best to weight sample data ... 200
Formula to be used in computing a general price index ... 201
Variable weights are desirable ... 202
Substitution of new commodities for old ... 203
Two separate indices of purchasing power desirable—trend index and cycle index ... 203
The rent of houses should appear in the cyclical index only ... 204
The price of labor ... 205
Ought bond prices to be considered in measuring the general price level? ... 206
Prices of stocks ought to be included in the cyclical index only ... 207
Feasible to compute a satisfactory general price-index ... 208
Commodities to be included in each type of index ... 208
Combining the cycle index with the trend index ... 209
The best standard of deferred payments ... 209
Index of price of direct goods probably best standard of deferred payments ... 213
A general deflating index not obtainable ... 214
Summary ... 216

CHAPTER IX

PAGE

CONCLUSION 219
The first purpose of this book..... 219
The second purpose of this book..... 219
The third purpose of this book..... 220

INDEX 221

INDEX NUMBERS ELUCIDATED

INDEX NUMBERS ELUCIDATED

CHAPTER I

INTRODUCTION

THE goal of this study is a most ambitious one. It is, in fact, nothing less than to dissipate, by the application of simple logic and careful reasoning, the mathematical fog which, almost from the beginning, has enveloped the science of index numbers. This is far from saying that the painstaking work that has been done in this field in the past is largely to be relegated to the junk heap. On the contrary, very little of it will be wasted. Existing studies along this line cover much of the ground satisfactorily. They discuss admirably most of the index numbers in common use, describing their peculiarities and showing in what ways the various series are similar and in what ways different. They analyze in great detail the characteristics of the various algebraic formulæ commonly used in the construction of indices. In all these respects, there is presumably little to add.

However, the fact remains that the leading writers in this field differ from one another radically concerning many of the fundamentals. Most of the disagreements are probably due to the same reason that furnishes the basis for the majority of disputes between scientists—failure to define

accurately the problem to be solved. It is, of course, imperative that the question to be answered be made absolutely clear-cut and specific—a rule which is basic in every scientific problem. Once the definition is made precise, the mode of solution often becomes so simple that one can scarcely conjure up again the difficulties which have long prevented progress.

A few eminent writers on index numbers have been prevented from arriving at sound conclusions by their failure to recognize the fact that the correct choice of an index number is dependent wholly upon the question for which an answer is sought. On the other hand, the necessity of having different index numbers for different purposes has been duly emphasized by a number of authorities.

The specific difficulty which seems most commonly to have prevented discussions in this field from bringing out with sufficient clarity the basic principles involved appears, however, to have arisen from the almost universal failure to concentrate attention sufficiently upon the fact that, since index numbers depend wholly upon sample data, the entire problem is one of the relation of the average of the sample to the average of the totality. As soon as this truth is fully recognized, it inevitably follows that most of the commonly accepted tests for index numbers go by the board, the whole question of the merits and demerits of index numbers is transferred to a new plane, and a considerable proportion of the discussion found in existing treatises is discovered to be entirely irrelevant.

According to the thesis set forth in this book the problem of index numbers always consists of two parts:

1. The determination of the exact mathematical formula which, when used with complete data, will answer the question asked. This is purely a matter of logic.
2. The discovery of the formula which, when applied to

a sample, will give results approximating as closely as possible those which would be given by the first formula if all the data were available and were used.

If this theory is correct, the merits of any proposed formula for a given index number can best be tested experimentally. So far as is known, no previous investigator has done any serious experimental work along this line.[1]

Chapters II, III, IV, and VIII are theoretical in nature and are devoted to explaining the fundamental principles governing index numbers. The doctrines here propounded will doubtless impress the adherent of accepted theories as being extremely bizarre. The reader must, perforce, form his own opinion as to the validity of the present thesis.

Unless the theories advanced in the theoretical chapters are fundamentally sound, the experimental statistical work described in Chapters V, VI, and VII mostly represents labor lost. The author, of course, hopes that it constitutes an important contribution to our knowledge of the facts pertaining to index numbers.

Much of the experimental work is devoted to an attempt to ascertain the merits of certain common formulæ when these formulæ are used in the computation of index numbers designed to answer certain specific questions. The reader is warned that the general conclusions deduced from this experimental work must not be thought to apply either to the answers to different questions or to the answers to the same questions under conditions radically different from those prevailing during the period studied.

[1] Professors Wesley C. Mitchell, Irving Fisher, and Maurice Olivier deserve full credit for their extensive statistical inquiries into the characteristics of index numbers. Their investigations have, however, not been directed to the end here discussed.

BIBLIOGRAPHICAL NOTE

The historical development of index numbers is discussed briefly by Wesley C. Mitchell in *Bulletin 284* of the United States Bureau of Labor Statistics (1921), pp. 7–11, and is treated more fully by Maurice Olivier in *Les Nombres Indices de la Variation des Prix* (1927), Chapter II.

The fact that the construction of an index series ought to be governed wholly by the question to be answered is stressed in the following articles:

Wesley C. Mitchell, *Bulletin 284,* United States Bureau of Labor Statistics (1921), pp. 23–59.

Arthur L. Bowley, *Economic Journal,* March, 1923, pp. 90–94.

G. Udny Yule, *Journal of the Royal Statistical Society,* May, 1923, pp. 424–430.

Carl Snyder, *American Economic Review,* September, 1924, pp. 416–421.

Donald R. Belcher and Harold M. Flinn, *Journal of the American Statistical Association,* September, 1923, pp. 94–97.

George H. Knibbs, *Journal of the American Statistical Association,* March, 1924, pp. 42–60; June, 1924, pp. 196–203.

F. Y. Edgeworth, *Economic Journal,* September, 1925, pp. 379–388.

Robert W. Burgess, *Journal of the American Statistical Association;* Proceedings, March, 1928, pp. 79–82.

That index numbers are samples and must always be considered on that basis was brought out by A. L. Bowley in *The Economic Journal,* March, 1923, pp. 90–94 and was emphasized by F. Y. Edgeworth in *The Economic Journal,* September, 1923, pp. 343–351, and in *The Journal of the Royal Statistical Society,* July, 1925, pp. 557–575. Correa M. Walsh takes issue with this view in *The Quarterly Journal of Economics,* May, 1924, pp. 500–519.

CHAPTER II

RELATIVE COMPARISONS USING COMPLETE DATA

More interest in relative than in absolute comparisons.—More people are interested in relative than in absolute comparisons, for the former commonly seem more significant: thus an increase of one milligram in the weight of a flea might be a matter of great scientific interest to the zoölogist, while the same increase in the weight of an elephant could never be discovered; there is more call for information concerning death rates than for the total numbers of deaths; for percentages of population increase than for the actual increases in the numbers of persons.

The conversion of absolute numbers to relatives.—What one does when one makes relative comparisons is to magnify the small items or minimize the large ones so that, at a selected time or place, called the *base,* all appear to be of equal size. Another way of expressing the same idea is to say that every item is expressed as a ratio to, or a percentage of, the size of the basic item.

To illustrate: In 1910, a certain hamlet had a population of 200, and a neighboring city contained 200,000 people. By 1920, the hamlet had added 40 and the city 20,000 new inhabitants, so that the new populations were 240 and 220,000 respectively. If we multiply the figures for the hamlet by 1,000 we get the following relative comparisons:

	Hamlet	City
1910	200,000	200,000
1920	240,000	220,000

Clearly the hamlet has proportionately grown faster than the city.

The process may be reversed and the city population divided by 1,000 giving the following comparison:

	Hamlet	City
1910	200	200
1920	240	220

The answer shows the same relationship.

A commoner way of making the comparison is to use the populations in 1910 as divisors. We then have

	Hamlet	City
1910	1.00	1.00
1920	1.20	1.10

The above comparison shows the *ratio* of the population in 1920 to that in the *base* year 1910.

Very frequently it is deemed preferable to express the comparisons in percentages. To do this the ratios are multiplied by 100. The results then appear thus:

	Hamlet	City
1910	100	100
1920	120	110

Each of the four relative comparisons given above has attained the same end—namely, to give to the hamlet and the city equal importance in the *base* year, 1910. In all four, the figures for 1920, whether expressed in terms of ratios, percentages, or otherwise, may be classed under the common name *relatives*. The term *index number* is often used as a synonym for *relative*. While, in general, there is perhaps no objection to such a usage, it is necessary, in this study, in order to avoid confusion, to confine the term *index number* to its strict meaning of indicator. In the following pages, therefore, the term *index number* is applied only to one special variety of the genus *relative*. Its nature will be discussed in the next chapter.

The advantage of relatives over absolute numbers.—The practical need for using relatives rather than absolute numbers is further illustrated by the following example:

If a statistician presents to his employer, a cotton manufacturer, the following figures:

	January 1	December 31
Price of cotton per pound (cents)	18.26	23.54
Average rate per hour for labor hired (cents)	31.32	38.17
Price of finished cloth per yard (cents)	26.14	34.82

the manufacturer will, as a rule, be unable to determine by looking at the data whether or not the selling price of his product has kept pace with the rise in material and labor costs occurring during the year. If however, the statistician takes the prices of January 1st as bases, dividing thereby the amounts for the other dates, the figures become:

	January 1	December 31
Price of cotton per pound	1.00	1.29
Average rate per hour for labor hired	1.00	1.22
Price of finished cloth per yard	1.00	1.33

The factory owner now sees at a glance that he has been able to increase the price of his product proportionally more than the rise in the cost of either labor or raw cotton.

The formulation of mathematical answers to questions relating to degree of change.—The only logical reason for taking the trouble to reduce numbers to a relative basis is to facilitate the answering of certain questions. A question is susceptible of statistical or quantitative treatment only if its answer can be formulated in algebraic terms. A corollary to this proposition is that an essential prerequisite to a comparison in the form of relative numbers is that the question to be answered be stated with precision. If the query itself is vague, obviously no one can say definitely what process should be followed in searching for the answer. It

is impossible, for example, to devise a formula which will correctly answer the question: *By what percentage did the population of New York City increase between 1880 and 1920?*

Between the specified dates the city expanded and took in areas formerly comprised in other cities. Are we to compare the populations of the different areas or of identical areas? If the latter, shall we use the area as it was in 1880 or in 1920? The Census of 1880 was taken as of June 1st and that of 1920 as of January 1st. Do we desire the population estimate for the same days of the month in the two years or are the Census dates the ones to be compared? Ought the immigrants at Ellis Island to be included in the city's population? What about those awaiting deportation?

Such queries as these must be definitely answered before the answer can be put into mathematical terms—before relatives for the two dates can be computed or the percentage of growth calculated.

For many executives, the exact formulation of a question is a difficult process, hence, in constructing relatives or index numbers, the statisticians employed upon an investigation frequently are forced to frame both questions and answers. When this is done, it is highly important that the reader should always be informed of the exact nature of the query to which the relatives give the correct answer.

The base defined.—The commonest practice is to make every relative number in the base year 100, in which case the relative number for any other time or place is evidently expressed as a percentage of the number chosen as a base. When treated as percentages, the data pertaining to the cotton mill previously mentioned appear as follows:

	January 1	December 31
Price of cotton per pound	100	129
Average rate per hour for labor hired	100	122
Price of finished cloth per yard	100	133

There is no imperative reason, however, for using 100 as the base number. Sometimes 1,000 is used. For many purposes it is decidedly best to follow the method first illustrated and use 1 as the base number, expressing all other numbers as ratios instead of per cents, the relatives being obtained by dividing all the items in the series by the base item. The use of 1 as a base greatly facilitates further mathematical manipulation of the numbers and hence is much to be preferred where such manipulation is applicable.

The process of comparison just outlined is applicable to a great variety of cases. Any number of series or items may be compared by using a common base for all. Take, for example, the following figures pertaining to the unfilled orders and earnings of The United States Steel Corporation for the years 1913–1918.[2]

TABLE 1

A COMPARISON WITH THE YEAR 1913 OF THE UNFILLED ORDERS AND EARNINGS OF THE UNITED STATES STEEL CORPORATION

YEAR	UNFILLED ORDERS AT END OF EACH MONTH		EARNINGS	
	Thousands of Long Tons	Relative, Base 1913	Thousands of Dollars	Relative, Base 1913
1913	5,907	100	11,432	100
1914	4,115	70	5,972	52
1915	5,189	88	10,866	95
1916	9,720	165	27,798	243
1917	10,716	181	24,608	215
1918	8,635	146	15,595	136

Clearly, in this instance, when all the facts are at hand, there is no difficulty in placing the yearly totals on a relative

2 United States Department of Commerce, *Survey of Current Business*, August, 1925.

basis by using each of the amounts in 1913 as 100. All that is necessary is to divide all the quantities in the series by the amount for 1913 and then multiply all the quotients by 100. This comparison shows at a glance that, proportionately, earnings fluctuated more violently than unfilled orders.

The usefulness of relatives not confined to time series.—One should not, however, fall into the error of supposing that relatives are useful only for comparisons in point of time. That such is not the case is illustrated by the comparisons in Table 2 showing the relative supply, in proportion to population, of both motor cars and income in various States.

TABLE 2

MOTOR VEHICLE OWNERSHIP AND INCOME PER CAPITA COMPARED FOR DIFFERENT STATES

STATE	ACTUAL NUMBERS OR AMOUNTS		RELATIVE TO CALIFORNIA TAKEN AS 1.00	
	Motor vehicles per thousand population,* 1921	Current income per capita,† 1921	Motor vehicles per thousand population, 1921	Current income per capita, 1921
California	193.5	$898	1.00	$1.00
Wyoming	138.2	749	.71	.83
Ohio	125.1	542	.65	.60
Vermont	105.7	515	.55	.57
Texas	100.2	433	.52	.48
Virginia	50.2	346	.31	.39
Alabama	35.0	250	.18	.28

* National Automobile Chamber of Commerce, *Facts and Figures of the Automobile Industry*, 1922, page 53.

† National Bureau of Economic Research, *Income in the Various States*, page 267.

All of the numbers in the right-hand section of the table are relatives using the figures for California as a base. The reduction of the data to the form of relatives makes much

more apparent the striking degree of correlation existing between income and automobile ownership.

The choice of a base.—In the example just cited, the State of California has been taken as a base. This means that it is put in the strategic position—that all other States are compared to it. If it is desired to lay the emphasis on Texas instead, then the ratios or relative numbers should be computed by dividing all the items in a given column by the item for Texas. If the purpose is to stress no one State more than another, one of the best procedures is to divide all the items in a series by the arithmetic mean (or, in other words, the arithmetic average) of the series.

In Table 3 and Chart I, the relative merits of two different types of bases are compared.

TABLE 3

RELATIVES COMPUTED ON TWO DIFFERENT BASES

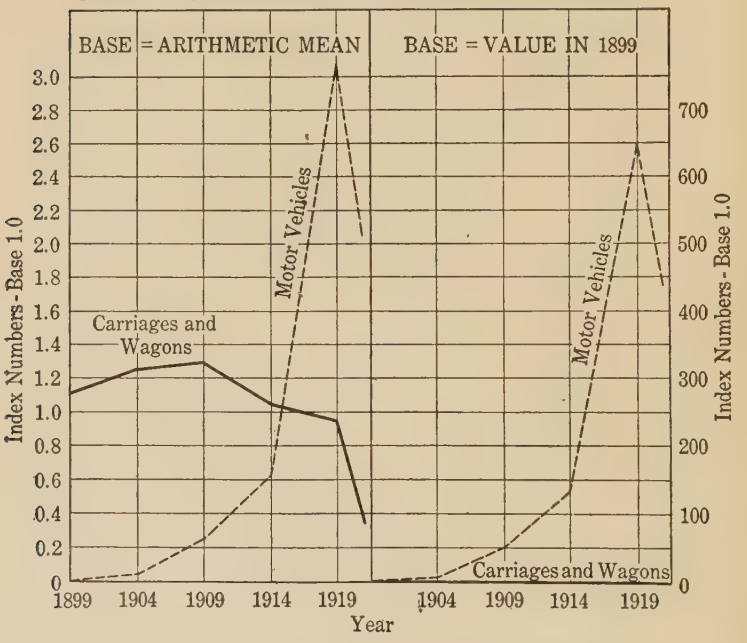
CENSUS YEAR	PRODUCTION * (000) OMITTED		RELATIVES			
			Base, Mean		Base, Value in 1899	
	Motor vehicles	Carriages and wagons	Motor vehicles	Carriages and wagons	Motor vehicles	Carriages and wagons
1899.....	$4,748	$138,262	.005	1.112	1.000	1.00
1904.....	30,004	155,869	.030	1.253	6.327	1.13
1909.....	249,202	159,893	.246	1.286	52.475	1.16
1914.....	632,831	131,547	.625	1.058	133.283	.95
1919.....	3,080,074	118,228	3.042	.951	646.719	.85
1921.....	2,079,404	42,446	2.053	.341	437.953	.31
Total.....	$6,076,263	$746,245	6.001	6.001		
Arithmetica mean.....	$1,012,710	124,374	1.00	1.000		

* United States Census of Manufactures, 1921, page 1035.

It is evident that the use of the arithmetic average as a base has certain marked advantages. One of these is that it keeps all the different variables as close together as possible, for, in each, the sum of the deviations above the base must equal the sum of those below. The mean is never a non-repre-

CHART I

EFFECT OF DIFFERENT BASES UPON GRAPHIC COMPARISONS BY MEANS OF RELATIVES



sentative base such as is the case with the production of motor vehicles in 1899. The use of the mean makes it easy to plot the relatives for motor vehicles and those for carriages and wagons in the same chart in such a way as to gain a clear idea of their respective trends. As the lower half of the chart shows, the case is very different when the

bases chosen are the values produced in 1899, for the output of motor vehicles in that year was far from typical for the period. As a result, when this base is used, it is impossible to compare the trends of the curves.

In addition, the use of the arithmetic average as a base furnishes a convenient means of verifying the computation of the ratios, for, if the computation is correct, the sum of the ratios must approximately equal the number of items in the series—a fact illustrated by the following figures:

Year	Tons	Ratio to Arithmetic Average
1922	24	0.92
1923	36	1.38
1924	32	1.23
1925	28	1.08
1926	20	0.77
1927	16	0.62
	6)156	6.00
	Average = 26	

On the other hand, this type of base has defects. One is that, whenever a new item is added to a series, it changes the size of the average—the number which is used as a base or divisor—hence all the relatives for that series must be recomputed or else it can no longer be said that the base is the mean of the series. Another objection to this base is that there is no date at which all the relatives equal 1.00 or 100, and hence it is difficult for the reader to locate the base merely by observing the figures. Furthermore, it is awkward, in writing an analysis of the tables, to describe the base both concisely and accurately—something that can readily be done when, in all series, the quantities at some given date are used as bases.

In the computation of price relatives, a cross between these two types of bases, namely the average for a given length of time, such as a decade, has often been used, the

object being to approximate the advantages of an average of all the years, while, at the same time, avoiding the necessity of recomputing all the relatives when later data become available. This plan undoubtedly has advantages, but against the advantages must be placed not only the weaknesses previously mentioned as inherent in the use of the average as a base, but also the fact that this device hinders the computation of ratios in every case in which figures are unavailable for even a single date in the base period. The longer the base-period the greater the probability that data will be missing for some part of the interval.

For the reasons mentioned, it is usually more convenient, and commonly just as satisfactory, to choose as a base a specific item, the size of which happens to be nearest to that of the arithmetic average, or better still, nearest to what the best evidence at hand indicates is for the series, the long-time mode or, in other words, the norm about which the items tend to vary. Such a base has most of the advantages of one covering a period of years and is a decidedly more convenient one with which to work.

There are cases, however, in which the use of a multi-period base is imperative. It has been found, for example, that no single year furnishes a satisfactory base for relatives designed for the purpose of comparing various crops either as regards yields or prices. The difficulty is that, no matter what year is selected for the base, it always happens that, in the season chosen, the yield of some important crop was either abnormally large or abnormally small and the price was, accordingly, far below or far above normal. When a series of relatives is based upon an abnormally high yield, it makes the yields of this crop for other years seem low in comparison to the yields of other crops. Similarly, if the price relatives for that crop are based upon a year in which the price was abnormally low, it will falsely appear that the given crop has been selling in other years at a price so high

as to be out of line with other crops. The reasons just cited have forced the United States Bureau of Agricultural Economics to use in its crop studies a base including a period of several years.

Many authorities are inclined to stress the importance of using the latest year as a base, the ground for their contention being that the majority of people are desirous of comparing all other dates with the present. This assumption remains to be proved. Is it not probable, indeed, that it is more difficult for most people to make comparisons backwards than forwards in point of time—in other words, is not the concept more easily grasped that prices have risen 66⅔ per cent since 1913 than that prices in 1913 were only 60 per cent as high as at present?

Without regard to which interpretation of popular psychology is correct, there are other objections to using the latest date as the base period. To do so requires that, each time the relative number or index is published, the figures for all earlier periods must be recomputed before they can be published. If publication occurs frequently, this necessity is likely to make the cost wholly prohibitive.

In practice, when a large number of items are being dealt with, to recompute on the new base all the relatives used is likely to take so long that, by the time the computation is completed, the last figure does not represent the present but rather a date well in the past.

Fortunately, in recent years, there has been a distinct tendency to base all kinds of historical index numbers upon the figures for the year 1913. This policy has much to commend it, not only because 1913 was a year nearly normal in many lines, but also because its use as a base has become so familiar to the public, that those interested in the course of some published ratios or percentages, fix in their minds the figures for important turning points and always have these ready for mental comparison with the latest figures

as they appear. There are, then, decided advantages in retaining this base for a number of years to come.

It is not feasible, at present, however, to base all relatives or index numbers on the 1913 figures, for, in many important fields, the collection of statistics began after that date. In such instances, the index-number maker must, perforce, use a later period as a base, and the wisest policy seems to be to choose for that purpose the latest year in which conditions appear to have been near the normal or trend,

If records for the great majority of series go back to 1913, it is, of course, feasible to insert relatives for the new items by using as the initial figure, in each instance, the average of the other relatives at that date.

All relatives using any fixed base necessarily share in a common weakness. It is impossible, without computation, to measure the percentage change occurring between two dates neither of which is the base. If the relative increases from 100 to 102, the 2 points rise marks a gain of 2 per cent, but if it rises from 200 to 202, the 2 points increase signifies a gain of but 1 per cent. Whenever it is desirable to measure the proportional change from the preceding year, the fixed base relative is usually found wanting.

It has been proposed to obviate this difficulty by constructing chain relatives in which every date for which figures are recorded is used as a base in computing the relative for the next period—in other words, the relative for any given date represents the ratio of the item for that date to the item immediately preceding.

Those advocating the use of chain relatives do so on the assumption that most of the persons interested desire to compare every period with the one immediately preceding. Evidence is wanting that such is the case. Apparently the demand for comparisons of this sort has never been strong enough to lead to the preparation and publication of such series covering any considerable period of time. Further-

more, it is almost impossible to put into words any question which the chain relative or index will answer.

It has been found easy to adopt in connection with a relative on a fixed base, an auxiliary device which meets the needs of those desiring to compare the relative for any period with that for the period immediately preceding. Such a comparison is secured by appending a separate column of figures showing the per cent of change since the preceding period. This device appears in the last column of Table 4 and is intended to be used in conjunction with the relatives employing the first year as a base. The table mentioned also gives a comparison of the results obtained by using the various bases just described with identical data.

TABLE 4

RELATIVE PRICES OF GRANULATED SUGAR AS ILLUSTRATED BY VARIOUS METHODS

(Price in the Base Year 100)

Year	Average Price per Pound*	Relatives on Following Bases				Per Cent of Change Since Year Preceding
		Average of first three years †	Last year	First year	Preceding year (chain links)	
1919	$.089	96	120	100		
1920	.127	137	172	143	143	+43
1921	.062	67	84	70	49	−51
1922	.059	64	80	66	95	− 5
1923	.084	91	114	94	142	+42
1924	.074	80	100	83	88	−12

* U. S. Bureau of Labor Statistics, *Bulletin 390, Wholesale Prices*, page 95.
† The base here equals $0.09267.

The device shown in the last column of the table, when presented alongside a relative using some fixed year as a base, possesses all of the advantages and few of the weak-

nesses of the series of chain relatives, and hence appears to offer a convenient method of satisfying readers having different interests.

Ratio charts.—Another device designed to accomplish much the same purpose has recently come into common use. Either the relatives are plotted on section paper having a logarithmic vertical scale or the relatives are first reduced to logarithms and these logarithms are then plotted on ordinary cross-section paper. A logarithmic curve obtained in either of these ways causes like *proportional* increments or decrements to appear as equal changes on the vertical scale. Such curves are excellent for comparing proportional changes either of different relatives or of the same relative at different dates. Of course they show little concerning absolute amounts and, since most readers are inclined to translate graphs into absolute quantities, they are somewhat more liable to misinterpretation than are curves plotted on the natural scale.

Relative comparisons representing a group of series.—Thus far, our attention has been directed to the treatment of relatives representing but a single series. Very commonly, however, it is desired to make relative comparisons of groups of series considered as a whole. Before such comparisons can be made, it is manifestly necessary to combine the correponding items in the separate series into aggregates or averages. This combination and comparison is usually effected by one of the two following methods.

1. The original items are summated and the totals are reduced to relatives by dividing by the base total. (For illustration of this method, see Tables 5 and 10.)
2. The original items in each series are reduced to relatives by dividing by the base item in that series. The relatives for the different series are then averaged. (This is the procedure followed in Tables 6 and 7.)

In either case, the individual series may be combined as they stand, or, on the other hand, each may be assigned a

weight representing its assumed relative importance in the combination. (For examples of weights applied to the original items, see Tables 8 and 9: of weights applied to relatives, see Tables 11 and 12.)

The two procedures, while in some respects distinct, are by no means unrelated; in fact, as will be shown in the following pages, it is possible to arrive at the same results by using either method. Without painstaking analysis, it is, however, not easy to see just how the two procedures are related. Such an analysis is presented in the pages immediately following. In studying this analysis especial attention should be directed to three facts:

1. That either procedure may commonly be substituted for the other procedure, provided a particular set of weights is applied to the individual series before combining them by the alternative method;
2. That, in the second procedure, a shift in the base used for the individual series in converting the absolute numbers to relatives, is usually equivalent to a change in the weights assigned to the respective series.
3. That a change in base or a change in weights commonly implies a change in the question asked and hence a change in the answer required.

The significance of these rather highly abstract statements will be more easy to perceive when the concrete examples which follow have been carefully studied.

Different relatives required to answer different questions.—It is impossible to state which is the correct method of computing a relative number until one knows the question which it is desired to answer. The following examples will make this point clear.

A New York educator is anxious to compare the quality of work done in five different High Schools of the State. To answer this question he secures the average grades in certain studies as shown by the results of the Regents' examinations. They are as indicated in Table 5.

TABLE 5

COMPARISON COUNTING ALL GRADE POINTS OF EQUAL IMPORTANCE

Subject	Average Grade in High School				
	A	B	C	D	E
Ancient History	89	90	86	84	85
Algebra	72	71	74	76	77
Latin	82	83	81	78	75
French	87	86	84	83	83
Physics	70	70	75	79	80
Total	400	400	400	400	400
Relatives	1.000	1.000	1.000	1.000	1.000

If the question asked is: "How do the aggregrate grades compare if all average grade points are considered equally important?" there can be but one answer: "The schools are all on a par"; for the totals of the averages are identical in all schools.

Some persons, however, will contend that it is illogical to assume that the grade points shown in Table 5 are all equally significant. They will point out that these figures consist of averages and not of totals of the grades of all the separate students in the school. Were we to compare totals of individual grades, quite different relationships might be shown, for the different subjects would not be taken by equal numbers of students. Furthermore, a point change in a high grade is counted as equally important with a point change in a low grade, and a point in a relatively easy subject like Ancient History is scored the same as a point in the more difficult studies of Latin and Physics. The Principal of School D, for example, notes the fact that the pupils under his direction are relatively strong in Algebra and Physics but comparatively weak in History and French. He also observes that, in all schools, the grades in the first two subjects tend to average lower than do the grades in the last

subjects mentioned. He therefore infers that it is easier for a student to add one per cent to his grade in History or French than to do the same in Algebra or Physics, and believes, therefore, that the comparison given by the totals in Table 5 fails to answer the precise question which he has in mind, namely: "If the average grades given in School D are taken as a standard and if all subjects are made to have equal weights in that school, how do the grades of students in School D compare with the grades of the students in other schools?"

This question can be answered exactly by taking the average grades in School D as a base in each instance. These base grades may each be considered unity. The mathematical formula necessary to answer the specific question asked is known as the *simple arithmetic average of relatives.* It is characterized by the assumption that the values of all the basic items are identical. The process of making the comparison by this method is shown in Table 6.

TABLE 6

A COMPARISON MAKING EACH SUBJECT OF EQUAL IMPORTANCE IN SCHOOL D

Method Used—Arithmetic Average of Relatives

Subject	Ratios of Grades in Various Schools to Grade in School D *				
	A	B	C	D	E
Ancient History	1.060	1.072	1.024	1.000	1.012
Algebra	.947	.934	.974	1.000	1.013
Latin	1.051	1.064	1.038	1.000	.962
French	1.048	1.036	1.012	1.000	1.000
Physics	.886	.886	.949	1.000	1.012
Total of above	4.992	4.992	4.997	5.000	4.999
Mean ratio; base 1	.9984	.9984	.9994	1.000	.9998
Mean ratio; base 1,000	998.4	998.4	999.4	1000.0	999.8

* Data from Table 5.

There can be no doubt that, according to the conditions of the comparison laid down by the Principal of School D, his pupils rank highest. Others may quarrel with his premises, but the mathematics of the solution is unassailable.

However, the Principal of School B may insist on comparing all schools on the basis of the grades in School B, although he agrees with the Principal of School D in everything except the school to be used as a base. Changing the base changes all the ratios and they now are as follows:

TABLE 7

A COMPARISON MAKING EACH SUBJECT OF EQUAL IMPORTANCE IN SCHOOL B

Method Used—Arithmetic Average of Relatives

Subject	Ratios of Grades in Various Schools to Grade in School B *				
	A	B	C	D	E
Ancient History	0.989	1.000	0.956	0.933	0.944
Algebra	1.014	1.000	1.042	1.071	1.084
Latin	.988	1.000	.976	.940	.904
French	1.0115	1.000	.977	.965	.965
Physics	1.000	1.000	1.071	1.128	1.143
Total of above	5.0025	5.000	5.022	5.037	5.040
Mean ratio; base 1.	1.0005	1.0000	1.0044	1.0074	1.0080
Mean ratio; base 1,000	1000.5	1000.0	1004.4	1007.4	1008.0

* Data from Table 5.

The insistence that School B be used as the standard or base produces results not at all gratifying to the pride of the students of that institution, for the mean ratios indicate that this change has forced their school definitely into last place.

The reason that the mean ratios in Table 7 differ from those in Table 6 is that the question asked by the Principal of School B differs slightly from the one asked by the Principal of School D. Its solution therefore requires a shift of the basis of comparison with the result, of course, that the answers arrived at are different. So much for the logic—what is the mathematical explanation of the difference in the mean ratios?

The immediate explanation is that the difference arises from a change in the weighting used. The shift of base to any given school increases the weight upon the subjects in which its grades are relatively low and lessens the weight upon those in which they are relatively high. In this instance, therefore, making School B the base places greater emphasis upon Algebra and Physics, the subjects in which the students of that school were weak. This change in emphasis is, then, responsible for ranking School B below School A.

A change in base equivalent to a change in weights.—That a change of base is equivalent to a change in weights is illustrated by the computations presented in Tables 8 and 9—tables which have been specifically designed to give the same mean results as those shown by Tables 6 and 7. In Table 7, the base has been changed from that used in Table 6; while, similarly, in Table 9, the weights are different from those used in Table 8. In both Tables 8 and 9, each grade recorded in Table 5 is assigned a weight such as to make the product in the base school equal 200. This product for the base year is chosen because it is ⅕ of 1,000 and, by its use, therefore, the total of the five subjects for the base school will automatically come to 1,000, a round number convenient for comparative purposes. The weights are, therefore, obtained in every instance by dividing 200 by the grade in the given subject in the base school. The weights in Table 8 are, then, obtained thus: $200 \div 84 =$

2.381, $200 \div 76 = 2.631$. Similarly in Table 9, $200 \div 90 = 2.222$, $200 \div 71 = 2.817$, and so on.

TABLE 8

A COMPARISON MAKING EACH SUBJECT OF EQUAL IMPORTANCE IN SCHOOL D

Method Used—Weighted Aggregate

Subject	Weight or Multiplier	Products Based upon Data in Table 5				
		A	B	C	D	E
Ancient History....	2.381	211.9	214.3	204.7	200.0	202.4
Algebra.....	2.631	189.5	186.9	194.7	200.0	202.6
Latin.....	2.564	210.2	212.8	207.7	200.0	192.3
French.....	2.409	209.6	207.2	202.4	200.0	200.0
Physics.....	2.531	177.2	177.2	189.9	200.0	202.5
Total of above (compare with Table 6).....		998.4	998.4	999.4	1000.0	999.8

TABLE 9

A COMPARISON MAKING EACH SUBJECT OF EQUAL IMPORTANCE IN SCHOOL B

Method Used—Weighted Aggregate

Subject	Weight or Multiplier	Products Based upon Data in Table 5				
		A	B	C	D	E
Ancient History....	2.222	197.8	200.0	191.1	186.7	188.9
Algebra.....	2.817	202.8	200.0	208.5	214.1	216.9
Latin.....	2.409	197.6	200.0	195.2	187.9	180.7
French.....	2.325	202.3	200.0	195.3	193.0	193.0
Physics.....	2.857	200.0	200.0	214.3	225.7	228.5
Total of above (compare with Table 7).....	.....	1000.5	1000.0	1004.4	1007.4	1008.0

It is clear now that the substitution of the grades of School B for those of School D as basic items resulted in increasing the multipliers for Algebra and Physics and diminishing those for all the other subjects. Evidently, then, the change in base has forced a change in weights. Experimentation will show that this is what always happens under like circumstances; hence we may lay down the general principle that, in making relative comparisons of the aggregates of a group of series, changing the bases is equivalent to changing the weights of the various items—a principle so important that the student of index numbers and their construction cannot well afford to overlook it.

With a common base the weighted aggregate may equal the simple arithmetic mean of relatives.—Because a slightly different procedure is used in arriving at the results, the relationship existing between a series of weighted aggregates and a series of simple arithmetic means of relatives is frequently overlooked. A comparison of Tables 6 and 8 or of Tables 7 and 9 will make it clear that the difference between these two types of series may be resolved into a question of weighting, for we see from Table 8 that, when weights are assigned to the different subjects which weights are of just the size necessary to make all products in the base school D equal in size, the aggregates of the products are identical with the mean ratios in Table 6 on the base 1,000. Similarly, the totals in Table 9 exactly correspond with the mean ratios in Table 7. It follows, then, that, by proper weighting, the method of aggregates may be made to yield a series of relatives exactly corresponding with the series obtained by averaging relatives.

The question asked governs the formula required for obtaining the answer.—Having digressed somewhat in order to consider the peculiarities of two different methods of making relative comparisons, let us now return to our analysis of the grades in the five schools which we have been consid-

ering. Perhaps, at this point, it may not be amiss to warn the reader against jumping to the conclusion that, after all, the various comparisons obtained by different mathematical processes represent different answers to the same question obtained by different methods of figure juggling, for such an assumption is absolutely contrary to fact. The idea that it is possible by applying different statistical methods to the same data to obtain various results having equal claims to validity is a pernicious but widely held fallacy which has done much injury to the reputation of statistical method as a science, and which, therefore, cannot be too strongly condemned. The truth is that the nature of the question asked determines absolutely the mathematical procedure which must be used in arriving at the answer, in other words, no essential change in the method of solution is permissible except when the question to be answered changes. In our illustration, for example, the change in base was due to a change in the whole premise concerning the proper way to measure education, the question answered in Table 6 pertaining to the relationship of the grades when School D is taken as a basis, while the query answered in Table 7 is how the grades compare when School B is chosen as the standard. This change in the question of course caused a variation in the method of computation and a shift in the results.

When different questions are asked, one must expect different answers. Thus when the schools are compared on the basis of giving equal values to all studies in School D, the results show that the work of Schools A and B are equally meritorious; but, when all subjects are given such weights as to give them equal importance in School B, the grades of School A appear to be slightly higher. Furthermore, on the first basis of comparison, D outranks E; but, on the second basis, E outranks D. Evidently, then, any alteration in a question, as, for instance, one involving a shift in the

basis of comparison, is likely to change the whole series of relatives constituting the answer to the question.

In our relative comparison of the grades in five schools, time was not a factor and a given school was used as a base. Perhaps, however, the most common type of relative comparison is one in which data for a certain period or point in time are used as the base, and all other quantities are related to those at the base time.

To illustrate this type of comparison, let us consider a question concerning the physical production of crops on the farms of a Southern county in which only two important crops are grown—namely, corn and cotton. Let us suppose that the production and value records are as follows:

	Production		Total Value	
	1919	1924	1919	1924
Cotton	2,400,000 lbs.	1,600,000 lbs.	$600,000	$320,000
Corn	240,000 bu.	480,000 bu.	240,000	360,000
			$840,000	$680,000

Suppose further that the question asked is: "Did the physical production of crops increase or diminish between 1919 and 1924?" What is the answer?

The total value evidently decreased, but, since prices per pound and per bushel both changed, this fact gives no clue.

If we convert production figures to relatives on the base 1919, we get

	1919	1924
Cotton	100	67
Corn	100	200
Average	100	133

These figures show an increase of one-third. But why are we justified in taking 1919 as a base—in other words in assuming that both crops were of equal importance in 1919?

If, instead, we assume that both were equally important in 1924 we arrive at the following relatives:

	1919	1924
Cotton	150	100
Corn	50	100
Average	100	100

This method of comparison leads to the conclusion that the total crop did not change in quantity.

The reason that the answers conflict is that the question is so vague that no definite answer can be given. Before an answer can be given, the query must be made absolutely specific as, for example:

"If a dollar's worth of each crop in 1924 is taken as one unit, how did the number of units produced in 1924 compare with the number produced in 1919?"

An answer beyond cavil can be given to the question in this form. It is:

	Size of Unit (Amount Selling for $1.00 in 1924)	Units Produced	
		1919	1924
Cotton	5 lbs.	480,000	320,000
Corn	$1\frac{1}{3}$ bu.	180,000	360,000
Total physical product	.....	660,000	680,000
Relative product; base 1924	.....	97	100
Relative product; base 1919	.....	100	103

But the question immediately arises as to why one who wishes to measure changes in physical production should take as units the amounts selling for $1.00 in 1924. Why not instead take the amount selling for $1.00 in 1919, or in some other year? Here we pass beyond the realm of the statistician. His duty is to answer the question asked. It is probable that no satisfactory reason can be assigned for

preferring as a unit the dollar's worth of 1924 instead of the dollar's worth of some other year; but a decision on this point must be made before the question becomes answerable. Once the question is made specific, the correct solution of the problem is no longer in doubt.

The principle that the statistician must know exactly what he is trying to find out before he can proceed intelligently, is so important that it perhaps warrants still another illustration.

Suppose, for example, that a question, which, on its face, appears perfectly definite, is asked: "Were average money wages in the Acme automobile factory higher or lower in 1927 than in 1919?" Let us assume that the plant employed but three classes of operatives, the rate of pay within each class being uniform. The daily rates were as follows:

	Daily Wage Rates	
	1919	1927
Foremen	$12	$12
Skilled operatives	8	8
Unskilled laborers	4	4

Apparently wages did not change. However, in 1919, the plant operated 300 days and, in 1927, only 240 days. For those who worked every day the plant was in operation, the following figures apply:

	Annual Earnings	
	1919	1927
Foremen	$3,600	$2,880
Skilled operatives	2,400	1,920
Unskilled laborers	1,200	960

Obviously there was a serious decline in the *annual* wages or *earnings* of each class.

On the other hand, the price of labor is usually measured by the rate per hour. Between the two dates, hours per day were reduced from 10 to 8; therefore hourly wage rates were as follows:

	Hourly Wage Rates	
	1919	1927
Foremen	$1.20	$1.50
Skilled operatives	.80	1.00
Unskilled laborers	.40	.50

Clearly, wages per unit of time rose in every occupation. New methods of operation, however, require a smaller proportion of skilled labor as is shown by the following figures:

	Hourly Wages; All Workers					
	1919			1927		
	Number employed	Rate	Total pay	Number employed	Rate	Total pay
Foremen	10	$1.20	$12	10	$1.50	$15
Skilled	50	.80	40	50	1.00	50
Unskilled	100	.40	40	500	.50	250
	160		$92	560		$315
Average rate per hour	$0.575			$0.5625		

Evidently there has been a slight decline in the average *rate per hour* paid to all workers.

What, then, is the answer to the original question? Did average money wages actually rise or fall? One cannot say, for the phrase "average money wages" has many meanings; hence the question is vague and indefinite. Once the query

is made specific, the answer also can be made definite and conclusive.

Relative aggregate values at various dates of a fixed number of shares of stock.—For our next inquiry, let us turn to a much studied field—namely the stock-market—and let us suppose that we are required to give a mathematical answer to the following question: "Had a man bought one share of each kind of stock sold on the New York Stock Exchange on January 2, 1924, what would have been the ratio of the value of his stocks at the close of the market on any specified day to the value of the same at the time of purchase?"

In answering this question, it is manifestly unnecessary to know for any company anything about the number of shares outstanding or the number of shares sold, for these items do not enter into the problem. It is possible, for example, that, on the given day, the sales of the most active stock may have totalled a thousand times as many shares as the sales of the least active stock. It may also have been true that some companies had a hundred times as many shares outstanding as did others. These differences, however, were of no consequence to the hypothetical purchaser just mentioned, for he bought but one share in each company.

In comparisons based upon aggregates of original items, the larger items dominate the results.—Since the purchaser bought one share, and one share only, of each stock, a brief consideration of the facts will make it evident that changes in the market prices of those shares which were valued highly at the date of purchase are likely to have a dominant effect upon the movements of the aggregate value of all the shares. For instance, a share of the highest priced stock may have cost him $400, while the cheapest share was purchased for only $1. In this case, 1 per cent gain in the value of the high priced stock would net him as many dollars profit as 400 per cent gain in the value of the lowest priced share.

One per cent fluctuations are very common; 400 per cent variations are extremely rare. For this reason, the relative changes in the total value of the stock purchased by the investor are likely to be due in the main to the fluctuations occurring in a few of the high priced shares, the absolute changes in the lower priced stocks being too small to affect the total materially. This state of affairs results from the fact that the bulk of his money was invested in the high priced shares, the inevitable outcome of buying one share of each stock.

The simple and direct method of computing the relative numbers desired is illustrated by the following example which assumes, for simplicity's sake, that only three stocks were listed on the New York Stock Exchange at the given date.

TABLE 10

RELATIVE NUMBERS SHOWING CHANGES IN VALUE TOTALS REPRESENTING ONE SHARE OF EACH KIND OF STOCK

Simplest Method of Aggregates (Relatives in Base Year Equal 100)

Stock	Price at Beginning of Year			
	1922	1923	1924	1925
Ajax Cotton	$40	$36	$24	$20
Chrome Steel	8	6	4	2
United Tea	400	440	480	520
Total	448	482	508	542
Relatives; base 1922	100.0	107.6	113.4	121.0
Relatives; base 1925	82.7	88.9	93.7	100.0

Evidently, the heavy decline in the value of the first two stocks did not prevent a sharp gain in the totals, for the price per share of United Tea was so great as to cause small relative changes in that stock to overwhelm the changes in the other two stocks.

Relationship between the simple aggregate of original data and the weighted arithmetic mean of relatives.—The illustrations in Tables 6 to 9 were used to bring out the relationships existing between relatives representing *weighted aggregates of original data* and *simple arithmetic averages of relatives.* We may now turn to an analysis of another relationship of the same sort, although in a certain sense opposite—namely that existing between a series of relatives

TABLE 11

RELATIVE STOCK PRICES ARRIVED AT BY THE METHOD OF WEIGHTED RELATIVES

(Price in Base Year = 100)

STOCK	WEIGHT	RELATIVE PRICE AT BEGINNING OF YEAR (Base Year for Individual Relatives, 1922)							
		1922		1923		1924		1925	
		Relative	Product	Relative	Product	Relative	Product	Relative	Product
Ajax Cotton.....	0.40	100	40	90	36	60	24	50	20
Chrome Steel.....	.08	100	8	75	6	50	4	25	2
United Tea.....	4.00	100	400	110	440	120	480	130	520
Total.....	4.48	.....	448	.....	482	.....	508	.....	542
Aggregate relatives; base 1922.....	.....	.....	100.0	.....	107.6	.....	113.4	.....	121.0
Aggregate relatives; base 1925.....	.....	.....	82.7	.....	88.9	.....	93.7	.....	100.0

derived from a *simple aggregate of original data* and a *weighted average of relatives.*

That the process shown in Table 10 (that of summating the original items and then reducing the totals to relatives) is equivalent to using individual relatives for the different stocks and weighting United Tea very heavily and Chrome Steel very lightly is made clear by the calculations in Table 11 which, when the same year is used as a base, yield results identical with those in Table 10.

If the base date for the relatives of the individual items is changed, the relatives for each stock are also changed, but, if the weights are varied just enough to compensate, the products, the totals, and hence, necessarily, the relative sizes of the totals remain unchanged. This case is illustrated in Table 12.

TABLE 12

THE EFFECT OF SHIFTING THE BASE WHILE RETAINING THE SAME PHYSICAL QUANTITIES

Comparison of Stock Prices by Method of Weighted Relatives

STOCK	WEIGHT	RELATIVE PRICE AT BEGINNING OF YEAR (Base Year for Individual Relatives, 1925)							
		1922		1923		1924		1925	
		Relative	Product	Relative	Product	Relative	Product	Relative	Product
Ajax Cotton	0.20	200	40	180	36	120	24	100	20
Chrome Steel	.02	400	8	300	6	200	4	100	2
United Tea	5.20	76.9	400	84.6	440	92.3	480	100	520
Total			448		482		508		542
Aggregate relative; base 1922			100.0		107.6		113.4		121.0
Aggregate relative; base 1925			82.7		88.9		93.7		100.0

It is important to note that, in this particular comparison, the shift in the weights used as multipliers has been so arranged that it does not connote a change in the physical quantities involved. These quantities remain constant, being in each case one share of each kind of stock.

Evidently this form of comparison permits, then, of changing the base of each of the relatives for the individual stocks without changing the relationship between the totals.

On the strength of the illustration just given, these principles may be laid down:

1. *It is possible, by selecting particular weights, to obtain the same relationships between the sums of the products of weighted relatives as those existing between the unweighted aggregates of original data.*
2. *A change in the base of the individual relatives will affect the relationships of the weighted aggregates of these items, unless the weights are so changed as to compensate exactly for the shift in base.*
3. *A shift in the base used in calculating relatives for individual items will not affect the relationships between the weighted totals for all the items when the weights are so varied that there is no change in the products, in other words no change in the relative physical quantities of the items included in the totals.*

Many index numbers have been considered bad or good because they are or are not affected by a shifting of the base of the relatives, but rarely if ever have the principles underlying this characteristic been fully elucidated. Apparently the basic principles involved are those set forth in the statements just made.

We have seen that, despite the peculiar weighting and method of calculation used in Tables 11 and 12, the final comparisons are identical with those recorded in Table 10. This identity of findings is due to the fact that the question asked has not been changed. Tables 11 and 12 merely represent circuitous methods of obtaining the same results arrived at simply and directly in Table 10.

Possible effect of using one base for prices and another for weights.—The United States Bureau of Labor Statistics, for some time, used the prices of 1913 as bases for all price relatives, but weighted these relatives by the estimated total value of the various commodities produced in 1919. It is perhaps worth while to note at this point the fact, that, *under certain conditions,* the use of one year as a price base and

of another year as the weight base gives rise to a definite mathematical relationship between the averages derived respectively by the two methods. This relationship may be stated as follows:

If no changes occur in the physical quantities of any of the commodities dealt in, and if the weights used are the total values of the respective commodities sold in the *specified* year, then it follows that the weighted averages derived by using as weights the sales-totals in the *base* year will, at all times except the base year, be SMALLER than the weighted averages derived by using as weights the sales-totals in any other year.

The reason for this is that, under the particular conditions specified, the use of any other year than the base year as a standard for weighting automatically gives weights heavier than those of the base year to the commodities that were higher priced in years other than the base year, and weights smaller than those of the base year to those commodities that were lower in price in years other than the base year.

The principle just cited, while of interest from a mathematical standpoint, is probably of little practical importance, for changes in the quantities sold usually accompany changes in price, and, under such conditions, the relationship stated does not necessarily hold.

In computing price-relatives, it is essential to compare the same quantities of identical commodities.—Comparisons of prices at different times or places are only useful when the various prices apply to articles of approximately the same type. It profits us nothing, for example, to know that a bushel of wheat is worth more today than a bushel of corn sold for last year. We are still in the dark concerning price changes even if we are told that No. 3 winter wheat sold last year for $2.10, while No. 1 Canadian Northern wheat sold yesterday for $2.20 a bushel. Erroneous comparisons of this type are, however, extremely common. It is not

unusual, for instance, to attempt to show the extent to which rents have risen by comparing the average rent per room at present with the average rent a generation ago—entirely ignoring the fact that rooms have changed in size, in the quality of their construction, in their furnishings, and in the amount of heat, light, and other services supplied.

Another error of this same type which is very common, but which is easily overlooked, arises from the comparison of the total values at different dates of various aggregates of commodities differing more or less radically in their composition. It is just as hopeless, for example, to gain any idea of the trend of commodity prices by knowing that 3 pounds of tea plus 4 pounds of soap sell for more today than 5 pounds of tea plus 6 pounds of soap did last year as it is to compare the price of one pound of soap last year and one pound of tea this year. Yet statisticians have frequently made just this kind of an error by using index numbers with changing weights to answer questions concerning changes in the total value of a specific group of articles. The cases are rare in which, in the computation of price-relatives, it is permissible to vary the quantities of the commodities used in the calculation; for the questions which people commonly ask, seldom call for price comparison of this type in order to answer them correctly. One or two important exceptions to this rule will be discussed in a later chapter.

Difficulty of comparing the prices of different budgets.—Because of the fact that the basic articles consumed by peoples far separated in point of time or space differ greatly from one another, it is not feasible to compute relatives which show accurately for such peoples the comparative prices of the direct or consumers' goods which they have used. It is, for example, extremely difficult to make valid comparisons of the present time and Colonial days in Massachusetts, or of the United States and China today, or even

of the farmer and the wealthy urban resident in the United States. The difficulty is that the articles which are important in one budget are either missing or unimportant in another. Rice, for instance, is the staff of life to the Chinese coolie, but is of no consequence in the diet of the average American. Anthracite coal constitutes a heavy item of expense to the present-day New Englander, but it was unknown to his Colonial forbears. The cramped but convenient apartment of the New York "cliff-dweller" has little in common with the roomy but ill-equipped farm house surrounded by spacious grounds.

Commodity substitution in price comparisons.—There is, however, an overwhelming demand for comparisons of the type just mentioned. Almost everyone is anxious to know how the purchasing power of the wage earner's dollar has changed—how much an ounce of gold will buy in Shanghai as compared to its purchasing power in New York. Presumably, the average inquirer has in the back of his mind the idea of comparing the cost in the two times or places of a certain quantity of gratification.

Although accurate measurements of this sort are unobtainable, it is, however, feasible to approximate very roughly the figures sought by finding ratios of the prices of different articles supposed to serve the same purpose equally well. It is not wholly absurd, for example, to compare with the price of a bushel of potatoes, in the United States, the price in China of the amount of rice containing the same number of calories of energy, for both rice and potatoes are similar in composition and play the same rôle in the diet. This process has, however, distinct limitations, for one cannot find in the typical farm community a form of entertainment the price of which can fairly be substituted for a ticket to grand opera.

Composite relatives.—In certain cases in which direct comparisons are out of the question, one can still obtain rea-

sonably defensible comparisons by splicing together a series of relative comparisons.

Knowledge, for example, that a kilowatt-hour of electricity in 1927 sells for more than a candle in 1827, tells us nothing whatever about the course of prices of lighting facilities during the century. Though it is manifestly impossible to answer the question: "How have prices of electricity varied during the last hundred years?" because electricity was not for sale a century ago, we may, nevertheless, be anxious to have an approximate answer to the query: "What relative changes have occurred during the last hundred years in the cost of the lighting facilities commonly used by the majority of the inhabitants of Chicago?" To answer this question it is necessary to splice together several series of relatives by a process like the following:

TABLE 13

USE OF SPLICED RELATIVES TO MEASURE CHANGES IN THE COST OF THE MOST COMMON FORM OF LIGHTING

YEAR	PRICES OF COMMODITIES				INDIVIDUAL RELATIVES				COMPOSITE RELATIVES (Product of Individual Relatives)
	Candles, per pound	Kerosene, per gallon	Gas, per thousand cubic feet	Electricity, per kilowatt-hour	Candles, per pound	Kerosene, per gallon	Gas, per thousand cubic feet	Electricity, per kilowatt-hour	
1840	$0.10	.....	.....	.....	1.00	.....	.....	.....	1.00
1850	.12	.....	.....	.....	1.20	.....	.....	.....	1.20
1860	.14	.....	.....	.....	1.40	.....	.....	.....	1.40
1870	.12	$0.30	.....	.....	1.20	1.00	.....	.....	1.20
1880	.....	.24	.....	.....	.....	.80	.....	.....	.96
1890	.....	.21	$1.50	.....	.....	.70	1.00	.....	.84
1900	.....	.....	1.35	.....	.....	.....	.90	.....	.76
1910	.....	.....	1.35	$0.15	.....	.....	.90	1.00	.76
1920	.....	.....	.....	.12	.....	.....	.....	.80	.60

The first step is to compute an individual relative for each commodity. The base for "Candles" is the price in the

first year quoted. The base for each other article is the price in the overlapping year when both it and the commodity for which it was substituted were in use. The composite relative is built up from the individual relatives by a series of multiplications. Thus to compute the composite relative for 1880, the relative for 1870 (namely 1.20) is multiplied by 0.80 (the relative for kerosene in 1880). Next, the final relative for 1890 is secured by multiplying 1.20 by 0.70. The other figures in the composite relative are similarly calculated.

Identical results may be obtained by using a slightly different method of computation. For example, after it has been ascertained that the relative price of candles in 1870 was 1.20 one may divide this figure by \$0.30, the price of kerosene, at that date. The quotient thus obtained, namely 4, is then applied as a multiplier to the prices of kerosene in succeeding years. By this process, we obtain the composite relatives for these later years. The relative so derived for 1890 is 4×0.21 or 0.84. This relative is divided by \$1.50, the price of gas in 1890, the resulting quotient being 0.56. Multiplying the prices of gas in the various years by 0.56 gives the composite relatives 0.84, 0.76, and 0.76 shown in the last column of Table 13. The relative for the last year is similarly arrived at by multiplying \$0.12 by the quotient obtained by dividing 0.76 by 0.15.

By the process indicated in the table, we arrive at figures which seem to justify the conclusion that the price of commonly used lighting facilities has had a downward trend since 1840. This information could not have been obtained by the study of the prices of any single commodity.

With complete data exact answers are possible.— Throughout this chapter it has been assumed that all the data needed to answer a question are obtainable. In such cases, if the question is stated with precision, it is usually possible to devise a formula which, when applied to the data,

will give an exact answer to the query. To obtain accurate relative comparions under such circumstances requires sound mathematical logic and computations that are free from error. No problem of estimation or approximation is involved. The comparisons can always be made as accurate as the data themselves. In this respect, there is a distinct difference between problems of relative comparison when complete data are available and problems of relative comparison made by means of sample data or index numbers. Discussion of problems of the latter type will be taken up in the chapters that follow.

Summary

1. This chapter deals only with cases in which complete data are available.

2. For comparative purposes, relative numbers are much superior to absolute numbers.

3. Each relative comparison answers a precise question.

4. It is impossible to give a definite mathematical or statistical answer to a vague question.

With complete data available, it is, however, commonly possible to answer specifically a definite and pertinent question.

5. It is never possible, by the application of two or more statistical formulæ to the same data, to arrive at conflicting answers to the same question, all of which answers have serious claims to validity.

It is, however, often possible, by applying different methods to the same data, to arrive at the same answer by different routes.

6. Any change in the question asked usually results in a change in the answer.

7. Comparison is greatly facilitated by a wise choice of base for the relatives.

8. Chain relatives have little value; but the plan of making a subsidiary comparison by recording the percentage of change from the preceding year is one which can often be utilized to advantage.

9. Ratio charts, or in other words charts in which the data are plotted on a logarithmic scale, are extremely helpful in many instances but are not of universal applicability.

10. Relative comparisons of complete data may cover but a single series of items or, on the other hand, may refer to a combination of several series.

11. Identical comparisons are obtainable by using either relatives based upon aggregates of the original items or aggregates of relatives calculated from the original items. The differences between these two methods may often be resolved into differences in weighting of the items.

12. A shift in the base used in converting original data to relatives is usually equivalent to a shift in weights, and commonly represents a change in the nature of the question for which an answer is sought.

13. To measure correctly changes in the average price of a group of commodities, it is usually essential that the composition of the group of commodities be the same at all dates compared. In some cases, however, it is permissible to substitute, by proper methods, a new commodity for one that has gone out of use.

QUESTIONS AND PROBLEMS

1. What is the essential difference between a series of relatives and a series of index numbers?

2. Cite instances in which absolute comparisons are more significant than relative; relative more significant than absolute.

3. From a given series of complete data, can you obtain one or many correct series of relatives?

4. If wheat production rose from 600 bushels in 1890 to 800 bushels in 1900, and to 1000 in 1920, did the earlier or the later period show the greater

(a) Absolute increase?
(b) Proportional increase?
(c) Rate of increase?
(d) Proportional rate of increase?

5. Would it be more convenient to deflate a series representing total money values by using as divisors a series of relatives on the base 1 or a series of relatives on the base 100?

6. Would you be more interested in comparing 1920 prices with 1919 prices than you would be in comparing 1920 prices with 1913 prices?

7. Were you called upon to choose a year between 1920 and 1925 as a general base for many series of relatives, which year would you select and why?

8. In choosing a base for time series, what objections are there to using for the purpose, in every case, an average of the most recent years?

9. Would the conversion of the absolute figures to logarithms facilitate the production into the future of the curve representing them?

(a) In a series in which the growth-rate is increasing? Why?
(b) In a series in which the growth-rate is slowing down? Why?

10. Write out several questions of the type that you feel ought to be answerable by a series of relatives. Now, in each case, write the mathematical formula which will give the relative answering the question. Can you write two formulæ applicable to the same question? If so, it is not definite.

11. Could you construct a good "cost of living" index for unskilled laborers if you had accurate records for each year of the average expenditures per laborer's family for food, for clothing, for housing, and for miscellaneous purposes?

BIBLIOGRAPHICAL NOTE

For a discussion of the advantages of relative over absolute comparisons, see:

Wesley C. Mitchell, *Journal of Political Economy,* July, 1916, pp. 669–671.

Irving Fisher, *Quarterly Publications of the American Statistical Association,* June, 1917.

James A. Field, *Journal of Political Economy,* October, 1917.

Robert E. Chaddock, *Principles and Methods of Statistics* (1925), pp. 433–439 and especially

Karl G. Karsten, *Charts and Graphs,* pp. 377–449.

As regards the choice of a base, it is worth while to read Robert E. Chaddock, *Principles and Methods of Statistics* (1925), pp. 189–193, and Oscar C. Stine, *Journal of the American Statistical Association,* Proceedings, March, 1928, pp. 92–94.

For a discussion of differences between averages of ratios and ratios of averages, see Irving Fisher, *The Purchasing Power of Money* (1911), pp. 198–225; G. H. Knibbs, *Prices, Price Indexes, and Cost of Living in Australia,* Commonwealth Bureau of Labor and *Statistics, Report No. 1,* December, 1912, Appendix VIII; Wesley C. Mitchell, *Bulletin 284,* United States Bureau of Labor Statistics (1921), pp. 68–81; Allyn A. Young, *Quarterly Journal of Economics,* February, 1923, pp. 342–364; Irving Fisher, *Quarterly Journal of Economics,* August, 1923, pp. 742–755; L. H. Bean and O. C. Stine, *Journal of the American Statistical Association,* March, 1924, pp. 30–35; and George R. Davies, *Journal of the American Statistical Association,* June, 1924, pp. 180–188.

The connection between the question asked and the formula required to compute the answer is covered in the references cited at the end of Chapter I.

The effect of using one year for the price-base and another year for the quantity-base is dealt with by Irving Fisher in *The Making of Index Numbers* (1922), Chapter III.

The difficulty of comparing budgets of different places or dates is well discussed by Arthur L. Bowley in *The Journal of the Royal Statistical Society,* May, 1919, pp. 343–361, and is touched upon by G. H. Knibbs in *Prices, Price Indexes, and Cost of Living in Australia,* Commonwealth Bureau of Labor and Statistics, *Report No. 1,* December, 1912, pp. 5–42.

CHAPTER III

THE NATURE OF INDEX NUMBERS

Why the subject of index numbers is ill understood.—At the present time, no device is more widely used by persons interested in the field of economics than is the index number. Many of the users, however, are not thoroughly familiar with the way in which index numbers are constructed, and hence their meaning is frequently ill understood. The non-mathematical person who takes the trouble to inquire into the nature of index numbers is too often discouraged, almost at the outset, by encountering a mass of difficult looking formulæ accompanied by descriptions either vague and indefinite or appallingly intricate. After an hour or two of study he is, therefore, likely to conclude that the subject is beyond his depth and discontinue his investigation.

It has well been said that the man who understands a subject well enough—even though the subject be extremely difficult—can elucidate it so that it is comprehensible to any intelligent person. Thus far, however, index numbers have rarely been explained in terms easily understood by the layman. Whether this failure to give an explanation both simple and lucid has been due to the fact that the writers themselves have not grasped the true significance of index numbers, or whether, instead, it has arisen from a feeling that simplification is unnecessary or beneath the dignity of a true scientist, the result has been to leave most students of the subject in the fog. It is hoped that the pages

which follow will present the matter in a form so definite that the principles underlying index numbers will no longer seem mysterious to any careful reader.

Index numbers a special class of relatives.—The last chapter treated of relative comparisons of *complete* data. In such cases, *index numbers,* as *strictly defined,* play no part. The solution of the problem was pictured as dividing itself into two parts:

1. Precise statement of the question and the mathematical formula required to obtain the answer.
2. The application of the formula to the data.

If the data are complete, the answer can always be found by this process. In practice, however, data are very often either incomplete or so superabundant in quantity as to make the expense of utilizing all the available figures prohibitive. In such instances, the statistician is compelled to substitute *estimates* for mathematical facts in attempting to answer the queries propounded. It is this necessity that has given rise to the invention of the device known as the index number.

As previously stated, the *index number* is merely a particular variety of *relative.* Certain writers have used the terms *relative* and *index number* interchangeably. Others appear to have differentiated *index numbers* from *relatives* on the ground that the latter are simple and homogeneous while the former are complex. Waiving the question of names, such a basis of classification seems to have nothing to commend it. *Relatives* may be either very simple or extremely complex. The population of a great city is usually treated as a single relative though it is composed of a thousand different groups, each varying in its own way. Relatives may, in fact, be much more complex than index numbers. The only differentiation, then, which seems worth

making is the segregation of relatives based upon *sample data* from those calculated from *complete data.* We shall arbitrarily confine the use of the term *index number* to the type of relative arrived at through approximations or estimates derived from *sample* data. This terminology will, in practice, be found to do little violence to established usage.

Most of the principles that have been discussed in connection with *relatives* in general apply with equal force to that special subdivision of relatives here referred to as *index numbers.* This statement should be emphasized in connection with the sections of this book dealing with the necessity of accuracy in formulating the questions to be answered, and with the definition of, and the choice of, a base—including the study of the results obtained by the use of different types of bases.

Index numbers defined.—In ordinary English, the word *index* means an indicator or something which stands for or represents something else. The name *index number,* therefore, suggests an agent and not a principal, and it is to such a usage that we shall here confine the term. In this sense, index numbers may be defined as relatives derived by some sort of an averaging process from sample data and intended for use either as indicators of the relative sizes of different groups of items or as indicators of changes occurring in aggregates during the passage of specified intervals of time. A comparison of such an index number with a series of relatives based upon complete data is shown in Table 14.

The figures in this table indicate that the sample used for the purpose of constructing an index of earnings constituted about one-seventh of all the workers in the establishment. As the sample was carefully selected, the differences between the average earnings per hour as shown by the sample and average earnings as revealed by the complete data are not great. Nevertheless, there are differences, and hence the index numbers are not identical with the relatives which

they are intended to represent. If the size of the sample were increased, the differences between the indices and the relatives would diminish, and if the increase were continued until the sample included all of the items, the index numbers and relatives would necessarily become identical. Relatives, then, are the limits which index numbers tend to approach as the samples used increase in size.

TABLE 14

THE RELATION OF INDEX NUMBERS TO THE RELATIVES DERIVED FROM COMPLETE EARNINGS DATA FOR A CERTAIN ESTABLISHMENT

YEAR	RELATIVES DERIVED FROM COMPLETE DATA				INDEX NUMBERS BASED UPON SAMPLE DATA			
	Wages paid to all employees (thousands)	Total hours worked by all employees (thousands)	Average earnings per hour	Relative hourly earnings, base 1905	Total wages paid to selected group of employees (thousands)	Total hours worked by selected group (thousands)	Average earnings per hour	Indices of hourly earnings, base 1905
1905	$919	4,238	$0.217	1.000	$135.9	612	$0.222	1.000
1909	1,329	5,879	.226	1.042	192.8	842	.229	1.032
1914	1,321	5,719	.231	1.065	190.3	810	.235	1.059
1919	2,586	5,432	.402	1.853	381.3	963	.396	1.784
1921	2,001	5,211	.384	1.770	308.5	795	.388	1.748

The perfect index number may, therefore, be defined as some average or combination of sample items which fluctuates relatively in exactly the same manner as does the series of relatives derived from the totals of which the sample is intended to be representative. In practice, of course, this perfect index number can never be attained—it is merely the ideal which the statistician strives to approach.

It may be well at this point to reiterate the fact that, when complete data are available, index numbers are only utilized when it is felt necessary to lessen the work of computation by using samples rather than all the figures at hand.

The problem of index numbers.—The problem of index numbers is nothing but a special case of the problem of sampling. A large sample, well selected, will doubtless yield an average which will fluctuate in much the same way as would the average of the totality of items (in other words of the complete data) were all the figures obtainable. A small sample, poorly chosen, is likely to give an average which varies in a manner radically different from the average of all the data.

Constructing index numbers is generally laborious work. The number of items available for sampling is often very small. The problem of the index maker is how to get satisfactory results from the materials at hand. In every case, there will exist certain limits within which errors are too small to be of consequence. The best index maker is the one who can so select his sample and so treat it mathematically as to use the smallest number of items that will keep the errors below the permissible limit.

Index numbers only useful to answer precise questions.—It goes without saying that one cannot hope to get accurate answers to questions from sample data when the queries are of such a nature that they could not be answered from complete information. No mathematical formula can be devised which will make possible a positive answer to a vague question. The need, therefore, of having questions stated precisely before attempting to answer them through the aid of index numbers cannot be emphasized too strongly.

Presumably a few illustrations will aid in clarifying this statement. Suppose, in a given city during a specified year, the facts were as follows: *Wage rates per hour rose in mercantile establishments but remained stationary in manufac-*

turing. However, a reduction in working time made wages, when figured by the day, show a decline in the manufacturing field. Because of increasing unemployment, the monthly earnings of all workers depending on the industry for support were stationary in the mercantile field but fell sharply in manufacturing. Let us suppose further that the question asked is: *Did wages rise or fall during the given year?* Clearly the question is too indefinite to permit of a specific answer. If, however, the question is revised to read: *What was the trend during the year in the average wage rate per hour for all employees in manufacturing and mercantile establishments combined?* it becomes possible to frame a formula, which, applied to complete and accurate data, will yield a series of relatives *exactly* solving the problem.

Still another example may prove helpful. A statistician of high standing once stated as a self-evident truth that, since two index numbers of wholesale prices differed, at least one must be in error. He illustrated his point by saying that, if one index showed that prices were higher in London than in New York, while another showed that, at the same time, prices were higher in New York, both manifestly could not be right.

As a matter of fact, each statement might be entirely accurate, for the question of where prices are higher is extremely vague. Woolen goods might be higher in New York and pork in London; brick in New York and lumber in London; and so with any other pair of commodities or any two different combinations of articles. An investigating commission once found that, when measured by the list of articles used by workers in Country A, the "cost of living" was lower in Country A than in Country B, but that, when the budget of the working classes in Country B was taken as a standard, the reverse was true. Clearly, under such circumstances, it was wholly impossible to say whether the "cost of living" was higher in Country A or in Country B. Until

the question was framed with greater precision, it could not be definitely answered.

As a third illustration of the principle involved, we may take the case of prices of corporation stocks. Let us assume that we have at hand a record of the price at which every share of stock sold at a given market on each of two different days separated by an interval of time. Let us further assume that the number of skilled computers available is unlimited. Finally let us imagine that the following questions are propounded:

1. *How did the average price of all shares dealt in on the last day compare with the average price on the first day?*
2. *How would the total value of stocks on the last day have compared with the total value on the first day if the same number of shares of each kind of stock had been sold on the last day as on the first?*
3. *Considering each stock as of equal importance, what was the average ratio of the price on the last day to the price on the first day?*
4. *What per cent would an investor have gained or lost through changes in stock values had he bought one share of each stock dealt in on the first day of the period and sold out on the last day of the period?*
5. *When all the stockholders on the first day under consideration are taken as a unit, what per cent of gain or loss in the value of their securities is shown by the prices on the last day?*

Different formulæ are required to answer different questions.—To answer the first question it is necessary to divide the total value of sales on each day by the number of shares sold on the same day, and then to find the ratio of the quotient for the second day to the quotient for the first day.

The answer to the second question is obtained by multiplying the price of each stock at both dates by the number of shares sold at the first date. The products are next summated for each date. The ratio of the sum for the last day to the sum for the first day is then found. This is the answer required.

Question 3 is answered by finding for each stock separately the ratio of its price on the second day to its price on the first day. The ratios for the second day are then averaged.

The answer to the fourth question is arrived at by adding together the prices per share of the various shares dealt in on the first day and then adding the prices per share that the identical list of stocks commanded respectively on the second day. The ratio of the second total to the first total is the figure asked for.

To reply to the fifth question, it is necessary to go outside the market to obtain additional information, for, in this case, one must impute a value to all stocks outstanding at the first date. To solve the problem, the procedure is to multiply the total number of shares of each stock outstanding at the first date by the price at each date. A sum of the products is then obtained for each date. The ratio of the sum at the last date to the sum at the first date answers the question.

The examples just cited seem sufficient to illustrate the importance of having the question to be answered stated precisely, for any one of the five answers might be given in reply to the question: *What was the ratio of stock prices at the last date to stock prices at the first date?* The point is that the question in this form is too vague to permit of a definite answer. The *conditio sine qua non* of all scientific work is, then, to state the problem unequivocally.

We see, therefore, that, although the same price data were used in every instance, a different formula was required to

answer each question. And these are only five questions out of a far greater number which might logically be asked. It is obvious, then, that the same totality of data may yield, in response to different questions, an unlimited number of relative comparisons, each differing radically from the others but each giving an exact answer to the question to which it applies.

If the ratios answering the five questions given in the preceding illustration happen to be nearly equal, it is a pure coincidence. They may differ widely. The fact that they differ does not reflect upon the accuracy of any one of them, for, if the data are complete and accurate, and no errors are made in computation, the answers obtained by the methods outlined must be correct. To seek for a general formula that will answer all types of questions is, then, evidently an absurd procedure. No one of the five formulæ is a whit better or a whit worse than the others. Each is exactly right in its place and wholly wrong in all others.

To combine the five formulæ or to average the five ratios in the hope of meeting the needs of the proponents of all five questions is a procedure which has little to commend it, for the results thus obtained represent a hodge-podge which it is beyond the power of the human mind to comprehend. The part of wisdom seems to be to select a single question and answer it as accurately as possible, while reserving the attempt to answer the other questions until there is time available to devote to their solution.

There can be no "best" formula for an index number.— It has been shown that, even with complete data, it is usually impossible to answer correctly different questions by aid of the same formula. The same principle governs just as truly when the data are incomplete or when but part of the data are used as an index of the whole. If there is an infinite number of series of relatives, all of which are correct, and if every series of index numbers represents nothing but an attempt to

approximate some one of these series of relatives, it follows that, in so far as the types of index numbers are successfully designed to fulfill their missions, there must be as many different "best" ones as there are different series of relatives answering different questions.

To state the same idea in other words: *The mission of the index number is to make the sample tell the same story as the totality.*

As has been shown, the mathematical formula required, and hence the series of relatives constituting the solution, is likely to change whenever there is the slightest variation in the question which the relatives are designed to answer. Let us suppose that, from a sample, a perfect series of index numbers has been derived, namely one in which the fluctuations correspond exactly to the fluctuations in the series of relatives derived from the totality. If, then, the formula applied to the totality of items (and hence necessarily the series of relatives derived by its use) changes, it is clear that unless there is likewise a change either in the sample or in the formula used in computing the index number, the fluctuations of the index will no longer correspond with those of the relatives obtained from the totality. Evidently, therefore, any change in the nature of the question asked is likely to call for the substitution of a new formula for computing the index number in order to assure the highest degree of accuracy in the index series.

If the correct answers to different questions, as derived from all the data, differ from one another, obviously the approximations to these answers obtained through the use of index numbers will also differ from one another and, whenever the index numbers approach perfection, will approximate the same degrees of difference found in the relative numbers derived from the totalities. The only possible defense for the belief that there is a single ideal all-purpose index number apparently, then, must be that all

index numbers are, at best, such crude approximations to the relatives which they are intended to represent that it is impossible to reproduce by their aid anything but the most outstanding characteristics of the answers to the questions propounded, the errors in the index numbers being so large as to overwhelm all minor differences. This view is, of course, exactly opposite to that now widely held by certain eminent statisticians, who contend that index numbers are rapidly becoming instruments of precision. In so far as these statisticians are right and index numbers are successful in reproducing the fluctuations of the totalities that they are intended to represent, the index numbers designed to answer different questions will differ from one another as widely as do the answers to the questions themselves.

This being true, it is no more reasonable to suppose that one index number can be found which will meet all demands in the best possible manner than it is to assume that a formula can be found which will solve all problems in arithmetic. The reverse of this proposition is that it is imperative to see that, in answering a question by aid of index numbers, the particular index chosen is computed according to the formula that gives the best results when applied to that particular type of question.

Supposed tests of the quality of index numbers.—The literature in this field abounds with discussion of various "tests" which supposedly have merit as criteria of the goodness or badness of formulæ used in index number construction. Among such tests may be mentioned the *circular* test, the *time reversal* test, the *commodity reversal* test, and the *factor reversal* test. Clearly, such tests are entirely inapplicable to formulæ for use with *complete data,* for such formulæ, depending as they do absolutely upon the nature of the question asked, are either right or wrong, and their correctness or incorrectness is determinable by the ordinary rules of logic. Furthermore, these supposed "tests" are no

more helpful in regard to *index numbers,* for it has been shown that an index number is good or bad just to the degree that it conforms in its movements to the course of relatives derived from complete data, and none of the "tests" mentioned throw any light upon the nature of this relationship. The statistician appears, then, to be fully justified in relegating to the realm of forgotten things not only the "tests" mentioned above but all others, which similarly do nothing but create puzzling problems of interpretation.

Summary

1. Most writers on index numbers have failed to explain these devices in terms simple enough to meet the needs of the average reader.

2. Index numbers, in the narrow sense, are relatives derived from sample data. The problem of index numbers is, then, a problem of sampling.

3. The test of merit for any formula is how closely the index series derived by its use approaches to the series of relatives calculated from complete data.

4. The larger and the more representative the sample used in the construction of the index number, the more likely it is that the index series will approach closely to the series of relatives derived from the totality.

5. Since it is impossible to give precise answers to vague questions when *complete* data are utilized, it is evidently equally impossible to accomplish this end by using index numbers derived from sample data.

6. There can be no "best" or "general purpose" formula which will give an index number suitable for answering all sorts of questions. As long as questions asked differ, the formulæ used must also differ, if satisfactory results are to be attained.

7. The commonly accepted "tests" for index numbers are entirely useless.

QUESTIONS AND PROBLEMS

1. Why do people use sample data to compute index numbers of stock prices on the New York Stock Exchange, when complete data are available in convenient form?

2. Would two statisticians, using the same formula and working with price records from the same market covering the same period, derive identical index series? Why or why not?

3. In what way is the problem of index numbers related to the problem of sampling?

4. Could you, by spending sufficient time and money, give a definite answer to the question: "Were prices higher in 1927 than in 1923?" Explain.

5. How would you proceed were you asked to construct "a general purpose index of prices"?

6. If two index series, representing wholesale prices in the same country at the same time, pursue distinctly different courses, would you conclude that one series is probably better than the other?

7. Is it certain that the formula which, with complete data, will give the precise answer desired is the best formula to apply to sample data in computing an index number?

8. Why is the search for an ideal index number formula necessarily futile?

9. What is the sole test of merit applicable to an index number?

10. Why is it fallacious to assume that the merit of a formula for index numbers can be judged by such criteria as the circular test or the time-reversal test?

11. Of questions answerable by a series of relatives, what question concerning the course of stock prices do you feel would be of interest to the largest number of investors?

BIBLIOGRAPHICAL NOTE

References covering the fact that the theory of index numbers is merely a special case of the theory of sampling and also the fact that the construction of the index must be dependent upon the question to be answered are found at the close of Chapter I.

The following sources give excellent descriptions of the methods used in constructing the more important of the index series published at the present time:

G. H. Knibbs, Australian Commonwealth Bureau of Labor and Statistics, Labor and Industrial Branch, *Report No. 1,* December, 1912, pp. 5–82*; Report No. 9,* July, 1919, pp. 230–235.

Wesley C. Mitchell, United States Bureau of Labor Statistics, *Bulletin 284* (1921), pp. 94–110; 175–336.

Irving Fisher, *The Making of Index Numbers* (1922), pp. 432–438.

United States Federal Reserve Board, *Federal Reserve Bulletin,* January, 1920, pp. 27–29; May, 1920, pp. 499–503; February, 1922, pp. 147–153; July, 1922, pp. 801–806; August, 1922, pp. 922–929; September, 1922, pp. 1052–1059.

Norman Crump, *Journal of the Royal Statistical Society,* March, 1924, pp. 183–193.

L. H. Bean and O. C. Stine, *Journal of the American Statistical Association,* March, 1924, pp. 30–35.

United States Bureau of Labor Statistics, *Monthly Labor Review,* July, 1924, pp. 98–99.

CHAPTER IV

SAMPLING AS RELATED TO INDEX NUMBERS

Complete data versus sample data.—As before stated, if complete data and sufficient clerical assistance are available, the derivation of the series of relatives necessary to answer exactly almost any precise query concerning comparisons is a simple matter. For example, if we wish to answer the question: *What relative changes have occurred in the average yield of wheat in the United States?* and if we have a record showing the total yield and the total acreage for every field in the United States in each year, we need only to summate the total yields, add up the acreages, and divide the total yield for a given year by the total acreage for the same year. The average yields per acre thus calculated can then be reduced to relatives on any base desired.

The process just described is simple and unequivocal, and, in the absence of errors in computation, the results obtained are indisputably correct. To secure data from every wheat grower in the United States is, however, a process so costly and difficult that, until recently, it has been essayed by the Bureau of the Census only once in a decade. The United States Department of Agriculture has, however, desired to obtain yearly comparisons. To accomplish this purpose, it has resorted to the use of sample data.

The problem of sampling.—The problem confronting it in this instance, may be taken to typify that which the maker of index numbers commonly must face, namely, that of

selecting from a large group or totality of items a relatively small group believed to be representative of the whole. Experience shows that, if the sample is large enough and properly selected, an average of the sample will be practically equal to an average of the totality. The statisticians of the Department of Agriculture are, then, justified in believing that it is possible, by proper methods, to derive from sample data approximately the same answer which could be obtained from complete data were they available. To accomplish this end four things are essential:

1. The sample items must be reasonably accurate.
2. The items selected for the sample must be sufficiently numerous to give approximate statistical regularity and to attain the degree of accuracy necessary in the case under consideration.
3. The method of selection must be such as to yield a fair or representative sample.
4. The formula used must be appropriate.

Let us consider these fundamentals *seriatim*.

Accuracy of sample items.—To obtain accurate estimates of yields per acre, the information must be secured from people who know the facts and are willing to tell the truth. The Director of the Division of Crop Estimates believes that the members of the large corps of voluntary *Crop Reporters* meet both these requirements. These *Crop Reporters* are usually farmers selected on the basis of intelligence and reliability, and it has been found that they are able not only to report accurately the acreage and output of the crop after it is harvested, but, indeed, that they can estimate the yield from the crop as it stands in the field, and do it with a surprising degree of accuracy. It appears, then, in this case, the sample items meet the requirement that they be reasonably accurate.

The most common purpose for which index numbers have

been employed in the past has been to portray changes in the averages of prices of one kind or another. One of the worries which has commonly beset makers of this class of index numbers has been the doubt that the price quotations obtained at different dates for nominally the same commodity really apply to identical articles. The cause for such doubt is accentuated in a time of inflation like that prevailing from 1917 to 1920, for, during this period, manufacturers were exceptionally given to changing sizes and grades while retaining the same name for the product. For instance a so-called pound can of fish shrank to 15½ ounces, then to 15 ounces, and later to 14½ ounces. A larger per cent of shoddy was introduced into a grade of woolen cloth passing under a certain trade name. A thousand similar instances could doubtless be cited. Such changes are likely to be overlooked by the reporter, untrained in statistics, who collects price data merely as a routine in which he is not interested. If half the changes are in each direction, the average may be but little affected, but, in a period when the tendency is all toward a lowering in weight or quality, any index number calculated from carelessly collected data is likely to be rendered materially inaccurate by the errors in the original quotations. To obtain thoroughly dependable price indices, it is, then, necessary to start with price quotations collected by enumerators who realize the necessity of quoting under the same title at different dates articles which are practically identical, and who are willing to expend the effort necessary to obtain such quotations.

Comparability at different dates is usually easier to attain if the successive quotations are secured from the same persons. When this is done, the informant comes to know for what grade quotations are desired, and the chance of error is therefore lessened.

Average prices.—When price quotations for individual articles are recorded as averages for a month or year, a

statement should always be given as to how the averages have been obtained. Do they represent the total value of sales divided by the total number of units sold? Have they instead been compiled by averaging the average quotations for the various days or perhaps the closing quotations for the different days? Have the quotations been secured from a single market or from a number of markets? In the latter case, have different weights been assigned to the various markets?

Clearly, in the case of a commodity in which there is a great variability in sales in each market, the various averages specified might show surprising differences in size. Each average may be ideal for answering some specific question, but be entirely unsuited for other purposes. Under these circumstances, it is evidently imperative that the nature of the average used be set forth clearly.

The number of items required for a satisfactory sample.—The statistician who sets out to answer a given question should determine at the outset what amount of error in the answer is too small to be of significance in solving the problem at hand. If, for example, an estimate of the average wheat yield, accurate to the nearest tenth of a bushel per acre, covers all needs, it is sheer waste of money to try to reduce the error to one-hundredth of a bushel per acre. It is well to remember that reducing the size of the error is likely to be an expensive process, for it can be shown mathematically that, under normal circumstances, the size of the error in the average varies inversely with the square-root of the number of items averaged. If, then, results accurate to one-tenth of a bushel are obtainable from a sample of 1,000 farms, it would require approximately 100,000 farms to reduce the error to one-hundredth of a bushel. Cost of collection of data and computation tends to vary in proportion to the size of the sample; hence, to divide the error by 10 multiplies the cost by 100. It follows that, as a general

rule, samples should be no larger than necessary to bring the results safely within the limits of error fixed by the necessities of the case. We cannot tell, however, whether or not the error has been brought within permissible limits unless we can approximate the absolute size of the error when the sample consists of some given number of items.

In practice, one of the most simple and effective ways of determining whether a sample is or is not large enough to give dependable results is to divide it into halves, each half being selected from the original data on the same basis,[1] and then to average separately the items in each half of the sample. The sample may then be halved again along different lines, and averages taken of the new halves. By repeating the process three times, six averages are obtained. The chances are that the divergence of the average of the sample from the average of the totality will not be greater than the difference between the largest and the smallest of the averages obtained from the sub-samples; hence, if this difference is less than the amount of error permissible in the final results, it justifies the assumption that the number of items is adequate.[2]

In cases in which the number of items in the totality is very large and the frequency distribution of the items therein is normal—namely is such as to give a bell-shaped curve when plotted—and when the method of selecting the sample is so designed as to prevent its containing any biased or cumulative error, it can be demonstrated mathematically that, in approximately nine cases out of ten, the difference between the average of the sample and the average of the totality

[1] If the separate items are recorded on cards, the simplest procedure is to shuffle the cards well, and then deal them alternately into two piles.

[2] As the errors of the parts of the sample are normally larger than the error of the entire sample, this rough-and-ready method may occasionally discredit a sample that is passably good. Since, however, the method errs, if at all, on the side of conservatism, it is not a bad one to use in practice.

will not be greater than the quantity obtained by dividing twice the average deviation of the items in the sample by the square root of the number of items in the sample. The average deviation of the items in the sample is usually arrived at by averaging the deviations of the respective items from the arithmetic average of the items in the sample. Equally satisfactory results can, however, be obtained by averaging the deviations from the mode or from the median.

This procedure is very simple, and, if the sample contains as many as 50 items, can usually be depended upon to show whether or not the number of items in the sample is adequate to produce results possessing the required standard of accuracy. Just as in the case when the method previously suggested is employed, the actual error is likely to be somewhat smaller than the one indicated, hence the use of this device is not likely to induce one to rely upon a sample too small to be dependable.

Random versus proportional sampling.—The third requirement specified above as being necessary for a satisfactory sample is that the items in the sample must be selected in an appropriate manner.

Two ways of sampling are in common use; first, the selection of items at random from among the totality of items, as illustrated by taking the first name from every page of a telephone book; second, the division of the totality into a number of classes and the selection at random within each class of a certain fraction of the items in that class, this fraction being roughly the same for all classes.

Sampling at *random* is not a method which, inherently, has any virtue. Every sample ought to contain as nearly as possible the same proportion of the items of each class entering into the totality. Frequently, the existing limitations of information and data make it impossible to secure samples that are even approximately proportional. The greater the extent, then, to which sampling can be done on the *proportional* basis, the greater the probability that it will

be truly representative of the totality. Random sampling as a complete method of procedure can be justified on no other ground than that it is an easy way to secure proportional sampling.

It is, however, usually necessary to use random sampling within the classes for which items have been selected on a proportional basis. It is permissible to use random sampling throughout with certain types of data. It is, for instance, appropriate when the totality of items to be sampled is large, and when the size of the items does not vary too greatly. If, for example, we wish to estimate the average heights of the men in a community, random sampling furnishes a convenient method of obtaining a correct average, provided, of course, that the number of items in the sample is not too small. On the other hand, a random sample of individual incomes might give an average very far from the truth, for the accidental intrusion or exclusion of a Rockefeller or a Ford would have a tremendous effect on the average of the sample.

Random sampling has the merit of eliminating any conscious or unconscious bias on the part of the person selecting the sample.

In the study of totalities having items varying greatly in size, it is essential to resort to *proportional* sampling. For the lower income classes, the United States Income Tax authorities can, for example, estimate with a fair degree of accuracy the amount of income derived from various sources by using as samples a certain fraction of the persons in each income class. However, in the classes containing the individuals receiving the larger incomes, the total number of individuals is so small that the sampling process is not appropriate. In such cases, reasonable accuracy can be attained only by utilizing all the items, thus avoiding the unnecessary errors which would grow out of the inclusion or exclusion of a few items of exceptional size.

In general, the United States Department of Agriculture,

in making its crop estimates, follows the method of proportional rather than the method of random sampling; for an effort is made to keep the ratio of the number of Crop Reporters to the total number of farmers about the same in all parts of the United States. Even though this procedure does not assure proportionality in the sample for the growers of any particular crop, the system works very well, for, in a State like South Carolina, where few farmers grow wheat, not many Crop Reporters will furnish estimates concerning the yield of this crop. In North Dakota, on the contrary, where wheat is the most common crop, yield estimates will be received from many Crop Reporters. In the case of wheat growers, the apportionment of the sample is, therefore, reasonably satisfactory.

The same is also true for other crops raised mainly by thousands of small growers. Difficulties arise, however, when the bulk of a crop is concentrated in a few counties. If, for example, the two or three Crop Reporters from Aroostook County, Maine, happened to be non-representative potato growers, the estimates for the whole New England potato crop might be sadly distorted, for Aroostook County produces more potatoes for market than are grown in all the rest of the North Atlantic area. In dealing with instances of this kind, the Department of Agriculture is, then, forced to take especial precautions to see that the leading growers in the most important districts are included among the Crop Reporters, for experience has shown that it is entirely unsafe to leave their selection to chance.

The instances just cited indicate that, despite the fact that the random method of sampling has the merit of eliminating any errors arising from personal bias, a carefully selected proportional sample is, as a rule, superior to a random sample, and that the extra work involved in apportioning the sample is usually justified except in those cases in which the range of the items in the totality is relatively small and

in which the sample consists of at least several hundred items.

The weighting of the sample items.—The fourth requisite —that the formula used must be appropriate—involves a great variety of considerations. The first topic that we shall discuss under this heading is the weighting of the items in the sample before they are combined in any form of average. When the items are not multiplied by factors before averaging, the average is said to be unweighted. The fact should be noted, however, that to say that the items are unweighted is equivalent to saying that they are weighted equally.

The weight assigned to an item always indicates the relative importance for the purpose at hand ascribed to that item. For example, in calculating the wheat crop of the United States, the Department of Agriculture weights the returns from the various geographical areas. The estimate for the United States is, then, not made up by throwing together the estimates of all the individual Crop Reporters. On the contrary, the total yield and the total acreage are estimated for each State and the State estimates are then combined to obtain a total for the nation. By this method, each State is weighted in proportion to the importance of its wheat crop and not in proportion either to the number of Crop Reporters in the State or to the amount of wheat raised by them.

From this illustration, we can draw a most important general rule which should be kept in mind by every maker of averages and index numbers. The rule is: *To secure correct averages, sample items must be weighted, not in proportion to the importance of the items in the sample, but, instead, in proportion to the importance of the sections of the totality represented by the respective sample items.* This general principle is so important that it may be well to illustrate it more fully. Let us suppose, for instance, that the statisticians in the Department of Agriculture

should discover that, on the farms of the Kansas Crop Reporters, only 20,000 acres of wheat were grown, while those of Nebraska Crop Reporters happened to include 30,000 acres of wheat. It would be entirely erroneous to assume that, in making up averages and relatives, Nebraska should be weighted 3 and Kansas only 2, for Kansas normally plants ten millions of acres and Nebraska only three millions. Evidently, sound statistical practice requires that such weights be applied to the items as will result in giving Kansas a wheat acreage 3⅓ times as large as that of Nebraska.

The method of weighted aggregates.—One way to attain the desired end is to multiply the sum of the acreages and also the sum of the yields in the Nebraska sample by the ratio of the total Nebraska wheat acreage to the wheat acreage of the Nebraska Crop Reporters (that is, by 3,000,000 ÷ 30,000 or in other words by 100) and likewise the sum of the items in the Kansas sample by the ratio of the total Kansas wheat acreage to the wheat acreage of Kansas Crop Reporters (that is, by 10,000,000 ÷ 20,000 or in other words by 500). This procedure gives aggregates approximating the total acreage and yields which would be obtained if the samples were expanded to include all the wheat growers in the two States. When the estimated total acreages of wheat all have been similarly computed for the forty-eight States, the results may be summated to give the total acreage for the United States. The total yield for the United States is likewise computed by adding the total yields for the separate States. The aggregate yield for the United States is then divided by the corresponding aggregate acreage in order to obtain, for the nation as a whole, the estimated average yield per acre. Averages thus calculated for various years are reduced to index numbers by dividing all of them by the average for the year which it is desired to use as a base.

A variation of this method which gives identical results

is to compute the average wheat yield per acre—as shown by the Nebraska sample—and multiply it by the estimated total Nebraska acreage (namely 3,000,000) to obtain the estimated total wheat yield for that State. Similarly, the average wheat yield per acre shown by the Kansas sample is multiplied by 10,000,000 to give the estimated total yield for Kansas. The total yields for the various States having been thus arrived at, the further steps necessary to calculate, for a series of years, index numbers of the average yield in the nation are identical with those described in the last paragraph.

The weighted arithmetic average of relatives.—It is possible to arrive at index numbers identical with those obtained by the *method of aggregates* described above, by using an entirely different procedure—namely, the following. The average yields per acre for a series of years in the same State are calculated from samples. These averages are reduced to relatives on the desired base by dividing all the average yields by the average yield in the year chosen as a base. The relatives thus obtained are all multiplied by the estimated total yield of wheat in the State in the base year. The process is repeated for the other States. The products for the forty-eight States in the same year are summated, and the sum thus obtained is divided by the sum of the forty-eight weights. The result is the index for that year. The index numbers for the other years are similarly calculated.

Advantages inhering in the average of relatives.—The procedure of calculating averages first and deriving all index numbers from these averages seems preferable in most instances, but it does not follow that it is the best to use in every case. It is, for example, not well adapted for use with small samples which change in composition from year to year. The difficulties involved in its application in such cases arise from the fact that the size of the items added may differ materially from the average of the items previ-

ously included in the sample or that items substituted may be very different from the ones displaced. The following numerical illustration will make this point clear.

TABLE 15

MISUSE OF METHOD OF AGGREGATES

CROP REPORTER	WHEAT ACREAGES AND YIELDS ON SAMPLE FARMS					
	1925		1926		1927	
	Acres	Total Yield	Acres	Total Yield	Acres	Total Yield
Caleb Brown	50	1,000	50	1,100		
Oscar Jones	100	1,300	100	1,510	100	1,700
James Smith	.....	.....	150	1,500	150	1,815
Ezra Wood	.....	.....	.....	.....	200	1,400
Totals	150	2,300	300	4,110	450	4,915.
Average yields	.....	15.3	.....	13.7	.....	10.9
Index numbers	.....	1.00	.....	0.90	.....	0.71

A study of the figures shows that, in every instance in which comparison is possible, the yield per acre was larger in 1926 than in 1925 and was still larger in 1927. Nevertheless, the average yields show a decline, the reason being that the farmers substituted for Caleb Brown had larger farms but secured much lower yields per acre. If Table 15 constituted a complete record of the wheat crop under consideration, the averages at the bottom would give a correct picture of the real situation, but they actually represent not complete data but merely records of four farms out of many thousands—and the rules applying to the totality of items are often entirely inapplicable to samples. Clearly, then, the method of aggregates, when applied to small changing

samples with varying weights, may yield index numbers which give anything but true pictures of the facts.

In the instance recorded in Table 15 the only possible inference from the evidence is that, from 1925 to 1927, the index numbers should rise. How can an index be constructed which will reveal this tendency? The most convenient method of dealing with cases of this kind is illustrated by the following table.

TABLE 16

INDEX NUMBERS CALCULATED FROM CHANGING SAMPLES BY MEANS OF PIECED RELATIVES

CROP REPORTER	AVERAGE AND RELATIVE WHEAT YIELDS ON SAMPLE FARMS					
	1925		1926		1927	
	Average yield	Relative	Average yield	Relative	Average yield	Relative
Caleb Brown	20.0	1.00	22.0	1.10		
Oscar Jones	13.0	1.00	15.1	1.16	17.0	1.31
James Smith	.....	.....	10.0	1.13	12.1	1.37
Ezra Wood	.....	.....	.....	.....	7.0	1.34
Simple average of relatives	.....	1.00	.....	1.13	.....	1.34

The introduction of new items into a sample composed of relatives.—The exact method by which new items are inserted into the sample is a matter so important as to demand careful attention at this point. The following general rule may be laid down: *Whenever it is necessary to introduce a new relative into a sample composed of relatives, the new relative should be made to equal, at the date of its introduction, the average of the other relatives in the sample*

at the same date. This rule may be followed in the construction of all varieties of index numbers consisting of averages of relatives. As a typical example of the application of this rule, we may take the way in which James Smith's record is introduced into the average. This is accomplished by making his relative number for 1926 equal to the average of 1.10 and 1.16, the relative numbers already in the sample, namely 1.13. The relatives for future years are computed by multiplying the actual wheat yields in the respective years by the ratio of 1.13 to 10.0, that is by $1.13 \div 10.0$ or 0.113.

The same procedure is followed in inserting the relative for Ezra Wood in 1927. This device enables one to substitute items indefinitely without doing violence to the facts.

What weights are appropriate?—It will be noted that, in this calculation, no weights have been used. Ought weighting to be employed? We are certainly not justified in using as weights the respective farm acreages of Brown, Jones, Smith, and Wood, for we have no reason to believe that, in the total area of which the sample is supposed to be representative, more of the wheat land is in 200-acre fields than in 50-acre fields. This is another instance of the fact that the relative importance of the items *in the sample* is usually of no significance whatever.

Granted that the farm areas shown in the sample are useless as weights, are there not other quantities which may legitimately be used for this purpose? Suppose, for instance, that, by reference to the Census or some other reliable source of information, we can ascertain the amount of the wheat area in fields of different sizes. Does this justify us in assigning these areas as weights to the yields of farms of corresponding sizes represented in the sample? This question cannot be answered without additional information. If it is true that the condition shown by the sample regarding the relationship between yield and size of farm is generally applicable—that is if large farms do, as a rule, have small

yields per acre, and vice versa, then we ought to weigh the farm of Ezra Wood by the acreage in large farms, the farm of Caleb Brown by the acreage in small farms, etc. If, however, so far as we know, there is no correlation between the size of the farm and the yield per acre, then we are not justified in using this basis in assigning weights to the items in the sample.

Some of the other peculiarities which must be considered in the weighting of sample items will be brought out later.

Continuous data may constitute misleading sample.—Table 15, shows how one may go astray by using averages of discontinuous samples as material for calculating indices. Strangely enough, a proposition which, on the surface, seems to be almost the reverse of the one just established also holds true, for the illustration recorded in Table 17 proves that one may fall into serious error by using as indices averages

TABLE 17

NON-REPRESENTATIVE AVERAGES OF RELATIVES OBTAINED FROM SAMPLE SALARIES IN THE CIVIL SERVICE

EIGHT INCUMBENTS IN FOUR POSITIONS	ACTUAL AND RELATIVE ANNUAL SALARIES							
	1923		1924		1925		1926	
	Salary	Relative	Salary	Relative	Salary	Relative	Salary	Relative
Betty Carr	$1,000	1.00	$1,040	1.04	$1,100	1.10		
Belle Dodd	.....	.....	.....	.....	900	1.07	$940	1.12
Mary Clark	1,400	1.00	1,450	1.04				
Madge Drew	.....	.....	1,200	1.03	1,260	1.08	1,320	1.13
Sumner Jones	2,400	1.00						
Thomas Reeves	2,000	1.00	2,040	1.02	2,100	1.05	2,140	1.07
Wm. Towns	1,800	1.00	1,840	1.02	1,900	1.05	1,950	1.08
Chas. Ware	.....	.....	.....	.....	.....	.....	1,800	1.10
Totals	$8,600	5.00	$7,570	5.15	$7,260	5.35	$8,150	5.50
Averages	1,720	1.00	1,514	1.03	1,452	1.07	1,630	1.10
Relatives of averages	1.00	.....	0.88	.....	.084	.....	0.95	

of relatives calculated from what at first blush seem to be ideal samples. To illustrate this statement, let us assume that our problem is to measure relative changes in the averages salaries paid to employees in the New York Civil Service. Let us further suppose that we select a number of representative positions and trace through successive years the salary records of their occupants. We shall arrive at the figures shown in Table 17.

In this case the relatives have been pieced in much the same way as in Table 16. For examples, Madge Drew's salary has been introduced in 1924 at 1.03, the average of the four other relatives in that year. Similarly, Belle Dodd's entry has been made in 1925 at 1.07, that being the average of the other four relatives at that date. By averaging the relatives thus obtained, we obtain index numbers purporting to show a steady rise in salaries during the four years. However, the indices entered on the last line—indices derived from the averages expressed in dollars—show a fall in salaries from 1923 to 1925 with partial recovery in 1926.

As a matter of fact, both sets of data are more or less misleading though both are helpful. The index numbers obtained by averaging the relatives lead one to believe that the trend of salaries in the New York Civil Service has been along a smooth curve, and that it has not fallen and then risen as shown by the salary averages in the sample.

The truth presumably is that the irregular course pursued by the arithmetic averages of the original items represents nothing but the result of chance fluctuations characteristic of such a small sample. The averages of relatives are then very helpful, for they reveal the fact that the trend is smooth rather than irregular. On the other hand, however, the averages of the original items are also enlightening for they indicate that the *upward* slope of the trend shown by the averages of the relatives is, in all probability, not at all

characteristic of the trend of the average wage of all New York civil employees. What has really happened is that men and women have been getting regular increases in salary throughout their respective terms of service, and it is these increases which give the upward trend to the individual relatives, and hence to the index number obtained by averaging them. However, since retiring employees are constantly being replaced with beginners drawing lower salaries, the increases arising from length of service do not necessarily give any upward tendency to the average pay of all persons employed in the Civil Service.

Many of the instances in which statisticians have arrived at different conclusions concerning the trend of wages, even though they were basing their reports upon the same data, have arisen from failure to take cognizance of the principles brought out by the above illustration. Since the same principles apply to other types of data also, the confusion from this source has, of course, extended to fields other than wages.

Summary

1. In selecting the samples to be used in computing index numbers it is necessary that—

   (a) The sample items be reasonably accurate.
   (b) The items be numerous enough to attain statistical regularity.
   (c) The items be representative.
   (d) The formula used in computing the index be appropriate.

2. The size of the error tends to vary inversely with the square root of the size of the sample.

3. Proportionality is the ultimate aim of all sampling. Random sampling is only one method used in seeking proportionality. Its only merits are that it is simple and that it avoids subjective bias.

4. Sample items should be weighted not in proportion to the importance of the items in the sample, but instead in proportion to the importance of the sections of the totality represented by the respective sample items.

5. When it is necessary to piece a sample, it is often better to obtain a series of index numbers by averaging the relatives than it is to derive the series from averages of the original data.

6. In introducing a new relative into a sample, care should be taken to see that, at the date of introduction, its size equals the average of the other relatives in the sample.

7. The mere fact that items in a sample are continuous from year to year does not prove that they are fair representatives of the totality.

QUESTIONS AND PROBLEMS

1. Why is the problem of index numbers to be considered merely as one phase of the problem of sampling?

2. Cite instances with which you are familiar in which the quality or size of a nominal unit of a commodity has varied although the name has remained unchanged.

3. If two samples are selected by the same method from the same totality, but if the respective index series derived from these samples differ widely in their behavior, how would you diagnose the difficulty, and what remedy would you prescribe?

4. Is it possible to secure a reasonably accurate index series from data which are seriously affected by:

(a) Compensating errors?
(b) Non-compensating errors?

5. Average the closing prices of all stocks sold on one day at the New York Stock Exchange. Now average the closing quotations for every tenth stock. Do the two averages differ materially?

6. If from a group of price relatives, all on the base 1913, two adequate samples are selected by applying the same rule, one sample covering the years 1913 to 1920 and the other the years 1921 to 1928, and if a separate index series is computed from each sample, would the two series thus derived, used singly or in combination, probably give a satisfactory measure of price changes between

(a) 1913 and 1920?
(b) 1913 and 1921?

(c) 1920 and 1928?
(d) 1920 and 1921?

Explain why or why not for each case.

7. What are the chief advantages of the random method of sampling? Of the proportional method of sampling?

8. Is the superiority of proportional sampling over random sampling likely to be more marked

(a) When the sample used is large or when it is small?
(b) When the items composing the totality vary greatly in size or when they are but slightly dispersed?

9. If you could obtain a correct statement of the income of every fiftieth family, as listed in a telephone directory containing 10,000 names, would this sample give you a fair representation of the incomes of the entire population of the city covered? Why or why not?

10. Suppose that, in a certain market, 1,000 commodities are dealt in, but ninety per cent of the total value of trade is confined to 80 commodities. If it is desired to construct a satisfactory index of the volume of trade or of prices in this market, is it probable that the index obtained from a sample of 400 commodities will come much closer to portraying the course of all commodities than will an index obtained from a sample consisting of the 80 most important commodities?

Do you feel that an index derived from a sample of 10 of the 80 important commodities would necessarily be more representative than another index derived from 10 of the much less important commodities?

11. Ought items in a sample to be weighted in proportion to their importance in the sample or in proportion to their supposed importance in the totality?

BIBLIOGRAPHICAL NOTE

Reference to those writers who have shown that the theory of index numbers is but a special case of the theory of sampling are given at the end of Chapter I.

The theory of sampling is discussed to some extent in most textbooks on statistics. The interested student will do well to read, in the order here given,

Harry Jerome, *Statistical Method* (1924), Chapter II.
Robert E. Chaddock, *Principles and Methods of Statistics* (1925), Chapter XI and pp. 383–387.
Frederick C. Mills, *Statistical Methods* (1924), Chapter XVI.
H. L. Rietz, *Handbook of Mathematical Statistics* (1924), Chapter V.

CHAPTER V

INDEX NUMBERS DERIVED FROM LINK RELATIVES OR RATIOS

Two different kinds of chain comparisons.—In an earlier chapter, the fact was noted that, for comparisons involving time, some authorities have favored the use of a series of index numbers, each giving the relative size of an item as compared with its size in the preceding period only. They have advocated this procedure on the ground that most people are interested primarily in comparisons with the last recorded date. In discussing this point of view, reasons were stated for believing that chain comparisons of this type do not meet the demands of the public and that all the advantages inhering in this method of presenting the facts can be obtained by using an index series having a definite base and showing, in a column parallel to the index series, the percentage of change occurring since the date immediately preceding. Chain comparisons by means of a series of relatives, each representing the ratio of the figure for the given date to the figure for the next earlier date, have, then, nothing whatever to commend them.

The frequently used type of index number derived from link relatives or ratios and referred to by many writers as a "chain index" has, however, little resemblance to the kind just described, for it does not purport to compare the size of a variable at a given date to its size at the date immediately preceding. On the contrary, the size of the variable

at all other dates is compared to its size at a single base date.

The advisability of using the term "chain index numbers" for a series giving comparisons with a fixed base is open to question,[1] and, to avoid confusion, we shall normally adhere to the more accurate, even though more awkward, terminology. We shall also keep in mind the fact that index numbers of this type are to be interpreted in exactly the same way as ordinary index numbers computed by the use of fixed-base relatives.

Use of link relatives forced by lack of continuous, comparable data.—The reason for using link relatives is not, commonly, that they are regarded as desirable in themselves, but rather because they are thrust upon the statistician. Unfortunately, he must work with such data as actually exist, and, all too frequently, it is impossible to secure strictly comparable numerical records covering the base date and all the other dates which it is desired to include in the comparison.

Suppose, for example, he wishes to construct an index of the price of automobiles. When he begins the collection of price data, he immediately discovers that most manufacturers change models every year. He finds, however, that, between July 1, 1923 and July 1, 1924, three leading models remained unchanged. Their prices were as follows:

	1923	1924
Speedwell six	$2,000	$1,800
Express, touring	1,000	950
Universal, roadster	700	700

By July 1, 1925, both of the first two cars had changed in type; hence, in comparing 1924 and 1925 prices, it is neces-

[1] Some authorities prefer the term "chain-derived index numbers." This title, though more nearly correct, is still doubtful as regards precision. Only the long designation—"index numbers derived from link relatives or ratios"—seems accurate, and it is decidedly awkward.

sary to substitute other cars. The comparable price quotations secured for this period are as follows:

	1924	1925
Universal, roadster	$700	$658
Crescent, sedan	1,100	1,023
Gibraltar, eight	3,600	3,600
Stanwix 40, coach	1,000	970

Clearly, any automobile price-index covering the period July 1, 1923 to July 1, 1925 must make use of the principle of substitution.

The problem raised by changing the composition of the sample.—When a substitution is made and the articles entering into the different links of the same relative are dissimilar, as, for example, when kerosene is substituted for candles, gas for kerosene, and electricity for gas, the danger always exists that the comparison thus obtained is unfair. Whenever an unfair relative is present in the group used in the construction of an index number, its presence is likely to vitiate the index number to a certain extent. The statistician who works with variable samples or who attempts to splice index numbers must always be on the alert or he will almost certainly fall into one of the numerous pitfalls which border such a course.

However, whenever we decide to use sample data, we are, consciously or unconsciously, pinning our faith to the laws of chance—are assuming that errors will tend to cancel each other. No matter how carefully it may be selected, the very fact that an item is a sample item indicates that it is only one of several items each of which might serve equally well as a criterion of the group which it is supposed to represent. Inherently, then, there appears to be no objection to substituting a new sample item for an old one, since, if both items are selected by following the same rule, there is no reason to believe that such a substitution is any more likely to increase than to diminish the difference between the average of the sample and the average of the totality.

Thus to return to the illustration cited above, it is just as probable that the changes in the prices of Crescent, Gibraltar, and Stanwix cars will represent the course of automobile prices in general as it is that the changes in the prices of Speedwell, Express, and Universal cars will perform that function. The problem of the index maker is to devise a mathematical process which will enable him to derive from a changing sample an index series which will give approximately as accurate an answer to the question asked as could be given were a continuous sample available. How can this be accomplished?

Linking the simple averages of relatives.—The method of handling the problem which, until recently, has been most commonly used, is first to average relatives for the separate periods, and then to link together the index numbers thus derived. This procedure may be illustrated by use of the automobile prices quoted above.

	Prices		Relatives	
	1923	1924	1923	1924
Speedwell six	$2,000	$1,800	1.00	0.90
Express, touring	1,000	950	1.00	0.95
Universal, roadster	700	700	1.00	1.00
Simple arithmetic average	.....	.....	1.00	0.95

	Prices		Relatives	
	1924	1925	1924	1925
Universal, roadster	$700	$658	1.00	0.94
Crescent, sedan	1,100	1,023	1.00	.93
Gibraltar eight	3,600	3,600	1.00	1.00
Stanwix 40, coach	1,000	970	1.00	.97
Simple arithmetic average	.....	.....	1.00	.96

We now have two overlapping pairs of indices and, by multiplying together the ratios 0.95 and 0.96, these may be linked together to form a single index series on the base 1923. The process is as follows:

	1923	1924	1925
Indices, base 1923	1.00	0.95	
Indices, base 1924		1.00	0.96
Pieced index series	1.00	0.95	0.912

Linking the weighted averages of relatives.—The simple arithmetic average of relatives gives equal weight to each car in the base year. Some questions could be better answered if appropriate weights were assigned to the different cars. When weighting is deemed desirable, all that is necessary is to multiply the original pairs of relatives by the weights chosen, summate the products for each date, and, in each instance, divide the sum of the products by the sum of the weights, thus:

	Weight	Relatives		Products	
		1923	1924	1923	1924
Speedwell six	1	1.00	0.90	1.00	0.90
Express, touring	4	1.00	0.95	4.00	3.80
Universal, roadster	3	1.00	1.00	3.00	3.00
Totals	8	.....	.....	8.00	7.70
Weighted arithmetic average		.....	.....	1.00	0.9625

For the later period, because of the introduction of new cars, it obviously may be necessary to assign new weights, even to cars appearing in the first period.

	Weight	Relatives		Products	
		1924	1925	1924	1925
Universal, roadster	2	1.00	0.94	2.00	1.88
Crescent, sedan	5	1.00	.93	5.00	4.65
Gibraltar eight	1	1.00	1.00	1.00	1.00
Stanwix 40, coach	3	1.00	.97	3.00	2.91
Totals	11	.....	.....	11.00	10.44
Weighted arithmetic average		.....	.....	1.00	0.9491

The final index is again obtained by multiplying together the averages of relatives.

	1923	1924	1925
Indices, base 1923	1.00	0.9625	
Indices, base 1924	....	1.00	0.9491
Pieced index series.	1.00	0.9625	0.9135

Average of relatives may be so weighted as to give results identical with those obtained by method of aggregates.—In recent years, for reasons which will appear later, there has been an increasing tendency on the part of makers of index numbers to rely upon ratios of aggregates rather than upon averages of relatives. In this connection, it should be remembered that there always exists a specific set of weights which, when applied to the relatives, will yield an index series identical with that obtainable by the method of aggregates.

It may be worth while, at this point, to illustrate this relationship by a numerical example. To simplify matters, we shall consider a case in which the sample is continuous throughout the period studied. Let us assume that we are trying to ascertain the relative levels at different dates of wage-rates for cooks in the cities of Ohio, and that, for this purpose, we have received continuous reports from typical employment agencies in a sample consisting of five leading

cities. We proceed by assigning to each city as a weight, not a figure recording the number of cooks in the given city but rather a figure representing the number of cooks in the area which the city is supposed to typify. In Table 18, for example, Zanesville is given a quota of 2,000, even though it contains far less than that number of cooks, the basis for its quota being that Zanesville is presumed to be representative of the smaller cities of the State. Since we are endeavoring to measure changes in the rate of pay of cooks of the same type in the same locations, the weights are kept constant.

To derive from a continuous sample the desired index numbers by using the method of aggregates, all that is necessary is to ascertain for each date the total pay for the 12,000 cooks on the basis of the numbers assigned respectively to each of the five reporting cities, and then to find the relative sizes of the totals as compared to the base. This process is illustrated in Table 18.

TABLE 18

INDEX NUMBERS DERIVED FROM AGGREGATES, THE LATTER BEING COMPUTED BY APPLYING CONSTANT WEIGHTS TO ORIGINAL ITEMS

(Base 1924)

City	Approximate Number of Cooks in Area City is Taken to Represent	1923		1924		1925	
		Weekly wage	Total weekly pay	Weekly wage	Total weekly pay	Weekly wage	Total weekly pay
Cincinnati	3,000	$20	$60,000	$20	$60,000	$22	$66,000
Columbus	2,000	22	44,000	21	42,000	24	48,000
Cleveland	3,000	24	72,000	25	75,000	26	78,000
Toledo	2,000	19	38,000	20	40,000	20	40,000
Zanesville	2,000	17	34,000	19	38,000	18	36,000
Aggregate pay	.....	.....	$248,000	.....	$255,000	.....	$268,000
Indices, base 1924	.....	.....	0.973	.....	1.000	.....	1.051

The reader will observe that, when the method of aggregates is used, as in Table 18, it is unnecessary to compute averages, for the index numbers can be calculated as readily from the totals as from averages derived therefrom. Thus, in Table 18, the index numbers are derived merely by dividing the aggregates by 255,000, the total for 1924, the base year.

Table 19 serves merely to illustrate the fact that the same index numbers may be obtained by using weighted arithmetic averages of relatives as by the method just described. The reader can, by experiment, readily establish the fact, that, whenever, as in this case, the original data treated by the two methods are identical, and when the weights chosen are such as to make the *products* in the base year also either identical, or in a fixed ratio throughout, the index numbers arrived at will always prove to be exactly the same. The system of weighting just mentioned is the one employed in Table 19.

TABLE 19

INDEX NUMBERS DERIVED FROM ARITHMETIC AVERAGES OF RELATIVES, WEIGHTS BEING KEPT CONSTANT

(Base 1924)

City	Weight	1923			1924			1925		
		Weekly wage	Relative	Product	Weekly wage	Relative	Product	Weekly wage	Relative	Product
Cincinnati..	60,000	$20	1.000	60,000	$20	1.000	60,000	$22	1.100	66,000
Columbus..	42,000	22	1.048	44,000	21	1.000	42,000	24	1.143	48,000
Cleveland..	75,000	24	.960	72,000	25	1.000	75,000	26	1.040	78,000
Toledo.....	40,000	19	.950	38,000	20	1.000	40,000	20	1.000	40,000
Zanesville..	38,000	17	.895	34,000	19	1.000	38,000	18	.947	36,000
Totals...	255,000	.....	.....	248,000	.....	.....	255,000	.....	.....	268,000
(Sum of products) ÷ (Sum of weights)...	.....	.....	.....	0.973	.....	.....	1.000	.....	.....	1.051

The following rule may, then, be laid down. *To make the weighted arithmetic average of relatives give results identical with those obtained by the method of aggregates, it is necessary to have the weights applied to the relatives proportional to the products in the base year, as obtained by the method of aggregates.*

Use of the method of aggregates with a changing sample.—Tables 18 and 19 have served to illustrate two ways of obtaining the same series of index numbers from the same continuous sample. But, as before stated, it is, in many cases, impracticable to secure a continuous sample. When the sample varies, how can the method be changed without destroying the comparability of the index numbers in the series? Suppose, for instance that, in the year 1926, the employment agency at Zanesville failed to report, but that data were secured from the previously non-reporting cities of Akron, Dayton, and Marion. How can this new material be used in constructing an index number by aid of the method of aggregates?

Index numbers derived by splicing ratios obtained by method of aggregates.—As previously stated, the best way to do this is to compute separate ratios for each series of links, and then to piece these ratios together to form a continuous chain. Clearly, when the chaining method is used, it is possible to introduce new items into the sample or to drop old ones out as often as may be necessary. In practice, it is seldom necessary to splice an index each year, as it is usually possible to secure a sample which will answer satisfactorily for several years.

In Table 20, which illustrates the process, the weights for the first three cities have remained unchanged, but it is not essential that weights be kept constant, except within the limits of a single link. The logic of the case determines what weights are to be used in each instance.

TABLE 20

COMPUTATION OF INDEX NUMBERS BY LINKING WEIGHTED AGGREGATES, KEEPING WEIGHTS CONSTANT WITHIN EACH LINK OF THE CHAIN

(Changing Sample)

City	Approximate Number of Cooks in Area City is Taken to Represent	1924		1925		1926	
		Weekly wage	Total weekly pay	Weekly wage	Total weekly pay	Weekly wage	Total weekly pay
Cincinnati	3,000	$20	$60,000	$22	$66,000		
Columbus	2,000	21	42,000	24	48,000		
Cleveland	3,000	25	75,000	26	78,000		
Toledo	2,000	20	40,000	20	40,000		
Zanesville	2,000	19	38,000	18	36,000		
The State	12,000	.....	$255,000	.....	$268,000		
Index (base 1924)	.....	.....	1.000	.....	1.051		
Cincinnati	3,000	.....	.....	$22	$66,000	$23	$69,000
Columbus	2,000	.....	.....	24	48,000	23	46,000
Cleveland	3,000	.....	.....	26	78,000	25	75,000
Toledo	1,000	.....	.....	20	20,000	22	22,000
Akron	1,000	.....	.....	18	18,000	19	19,000
Dayton	1,000	.....	.....	19	19,000	19	19,000
Marion	1,000	.....	.....	17	17,000	18	18,000
The State	12,000	.....	.....	.....	$266,000	.....	$268,000
Index (base 1925)	.....	.....	.....	.....	1.000	.....	1.008
Spliced index (base 1924)	.....	.....	1.000	.....	1.051	.....	1.060

It will be observed that the index numbers are computed directly from the aggregates, no averaging being necessary. A variation which may be made in the form of computation is as follows: First, multiply the aggregates in 1925 and 1926 by the ratio of the original aggregate for 1925 to the new aggregate for 1925 (*i.e.* by $\$268,000 \div \$266,000$). Next compute the index numbers for all years by dividing each of the adjusted aggregates by $255,000, the aggregate for the base year. This method will yield an index series identical with the one derived in Table 20.

TABLE 21

DEPARTURE OF INDEX SERIES DERIVED BY AVERAGING LINK RELATIVES FROM INDEX SERIES COMPUTED BY CHAINING AGGREGATES

	Commodity	Actual Prices per Unit						
		1921	1922	1923	1924	1925	1926	1927
METHOD OF AGGREGATES:	A	$48	$60	$64	$512	$1,024	$512	$48
	B	48	60	64	8	4	8	48
	C	48	60	64	128	512	128	48
	D	48	60	64	32	8	32	48
	E	48	36	32	64	256	64	48
	F	48	36	32	16	4	16	48
	G	48	36	32	128	256	128	48
	H	48	36	32	8	4	8	48
Total values		$384	$384	$384	$896	$2,068	$896	$384
Indices-fixed base.		1.00	1.00	1.00				
Relatives of aggregates		.....	.....	1.00	2.33			
					1.00	2.31		
						1.00	0.43	
							1.00	0.43
Spliced indices		1.00	1.00	1.00	2.33	5.38	2.33	1.00

	Commodity	Prices Relative to Those of Preceding Year						
		1921	1922	1923	1924	1925	1926	1927
METHOD OF AVERAGES OF LINK RELATIVES:	A	1.00	1.25	1·07	8.00	2·00	0.50	0.94
	B	1.00	1.25	1·07	0.12	0·50	2.00	6.00
	C	1.00	1.25	1·07	2.00	4·00	0.25	0.37
	D	1.00	1.25	1·07	0.50	0·25	4.00	1.50
	E	1.00	0.75	0·89	2.00	4·00	0.25	0.75
	F	1 00	0 75	0·89	0.50	0·25	4.00	3.00
	G	1.00	0.75	0·89	4.00	2·00	0.50	0.37
	H	1.00	0.75	0·89	0.25	0·50	2.00	6.00
Totals		8.00	8.00	7.84	17.37	13.50	13.50	18.93
Averages		1.00	1 00	0.97	2.18	1.69	1.69	2.37
Spliced indices		1.00	1.00	0.97	2.11	3.57	6.04	14.31

Characteristics of indices obtained by linking averages of relatives.—In the example illustrated by Table 20, we have used the method of aggregates. Now, since the method of aggregates corresponds to the arithmetic average of relatives with a particular weighting, the question naturally arises as to whether any peculiar virtue inheres in this specific form of weighting and whether, after all, one cannot place as much dependence upon index numbers obtained by averaging relatives as upon index numbers derived by the method of aggregates. To answer this query, Table 21 has been constructed. The hypothetical data contained therein are so arranged that, in each pair of years, the changes occurring illustrate a particular type of relationship existing between the two types of spliced indices—a relationship differing somewhat from that found in any other period.

It will be observed that the sample is a continuous one, and that, in the case of every commodity, the price is exactly the same in 1927, the closing date, as in 1921 the opening date. This arbitrary arrangement is necessary in order to demonstrate some of the principles which it is desired to elucidate. Let us now compare for each pair of years the two types of index series—namely, the chained ratios of aggregates and the chained averages of link relatives.[3]

1921–1922; equal fractional changes up and down.—In this case, manifestly, the total positive *percentage* changes exactly equal the told negative *percentage* changes. As it happens, in this particular instance, both methods give the same results. Were it not for the fact, however, that all of the items in 1921 are of the same size, the index numbers derived for 1922 would not be identical, a fact demonstrated by the following example:

[3] For many of the points noted in this discussion of the behavior of indices derived by chaining averages of link relatives, the writer is indebted to Dr. Frederick R. Macaulay of the National Bureau of Economic Research. Dr. Macaulay treats of the subject in the *American Economic Review*, March, 1916, p. 208.

	Commodity	1921	1922
Actual value	A	$48.	$60.
Actual value	B	16.	12.
		$64.	$72.
Relative of aggregates		1.00	1.12
Relative price of	A	1.00	1.25
Relative price of	B	1.00	.75
Average of relatives		1.00	1.00

We see that, when, as in the example just given, the *larger* number rises one-fourth and the *smaller* falls one-fourth, the relative of aggregates is *larger* than the average of relatives for the same year.

When, on the other hand, it is the *smaller* number that increases in size and the *larger* number that decreases, the relative of aggregates is *smaller* than the average of relatives —a relationship illustrated by the figures below.

	Commodity	1921	1922
Actual value	A	$48.	$36.
Actual value	B	16.	20.
		$64.	$56.
Relative of aggregates		1.00	0.88
Relative price of	A	1.00	0.75
Relative price of	B	1.00	1.25
Average of relatives		1.00	1.00

1922–1923; equal absolute changes up and down.—In this period, the *absolute* increases in price are exactly offset by decreases; hence the aggregates of absolute values remain unchanged. However, the $4.00 increase is a smaller frac-

tion of $60.00 than the $4.00 decrease is of $36.00. The result is that the average of relatives is forced downward while the relative of aggregates remains unchanged.

1923–1924; inverse proportional changes up and down in commodities having identical values.—In this interval, every *proportional* increase is offset by an inverse *proportional* decrease in the same ratio. Thus, the price of A is multiplied by 8, while the price of B is divided by 8; the price of C is doubled and the price of D is halved. Experience seems to indicate that this is the type of variation which commodity prices undergo when influenced purely by chance. In this particular instance, the relative of aggregates rises slightly more than does the average of relatives. Apparently, however, this slight difference is due merely to the accidental relationship existing in this illustration between the size of the sample and the size of the variable, for Dr. Frederick R. Maculay has shown that the normal tendency, under the circumstances, is for both indices to behave in the same manner.[4]

1924–1925; high-priced commodities increasing and low-priced commodities diminishing in value.—In this period, just as in the one immediately preceding, every proportional upward change was offset by a similar inverse proportional downward movement. Furthermore, the prices of all the commodities continued to move in the same direction as in the preceding period, that is those that were already high-priced continued upward and those that were low in price continued to go lower. The result of this tendency was to cause the average of relatives to rise much less rapidly than the relative of aggregates. This difference is due to the fact that the aggregates are dominated by the high-priced commodities, which are rising, while, in making up the aver-

[4] *American Economic Review*, March, 1916, p. 208.

age of relatives, each commodity is given equal weight, whether its value is large or small.

Since commodities that continue to rise in price eventually become high-priced, and those that continue to fall in price tend to become cheap, the following practical conclusion is indicated: *Whenever the condition of any market is such that commodities which have been falling in price generally continue to fall, and vice versa, it follows that an index series derived by chaining averages of link relatives will tend to have a declining trend as compared to an index series computed from aggregates.*

1925–1926; high-priced commodities falling and low-priced commodities rising.—Here we have an exact reversal of conditions in the preceding year—all the commodities that are high in price falling, and vice versa, the net result being to bring all the prices back to the exact level which they occupied in 1924. Again, every proportional increase was offset by an inverse proportional decrease, in the same ratio, applied to some other commodity. The relative of aggregates, being dominated by the action of the high-priced commodities, fell very sharply, while the average of link relatives, governed by the larger upward ratios due to the same fractional changes, rose very markedly, the net result being to make the final index based upon the latter nearly three times the size of that computed by the former method.

It will be noted that, although the price quotations in 1926 are identical with those in 1924, and the spliced indices based upon ratios of aggregates both stand at the same figure —2,333—the index series calculated from the averages of link relatives has nearly trebled in size, advancing from 2.11 to 6.04. This difference shows a radical weakness inherent in the latter method of computing an index series. This weakness arises from a mathematical quality of numbers illustrated by the following simple example.

	Commodity	1924	1925	1926
Actual prices	P	$12.	$36.	$12.
	Q	12.	4.	12.
	P + Q	$24.00	$40.00	$24.00
Relative of aggregates		1.00	1.67	1.00
Prices relative to preceding year	P		3.00	0.33
	Q		.33	3.00
Average of P and Q			1.67	1.67
Spliced index based upon average of relatives		1.00	1.67	2.78

When, in 1925, P is trebled the ratio is 3.00, or an addition of 2.00 to the base. When at the same time, Q is divided by three, bringing it back to its original position, the ratio is 0.33 or a subtraction of only 0.67 from the base—a quantity much smaller than the 2.00 added by the rise. The average of 3.00 and 0.33 gives a quantity greater than unity, indicating a rise in prices.

When the price movements are reversed, the decline in P gives a ratio of 0.33 and the rise in Q yields a ratio of 3.00. Again, the average of the ratios is 1.67 indicating a rise in prices.

Clearly, it makes no difference whether a number rises and then falls, or falls and then rises—in each case an increment is added to the base. We see, then, why it is that indices derived from averages of link relatives constantly tend to climb above the correct level whenever there is any tendency for prices to return to a normal after they have departed therefrom in either direction.

Another equally accurate way of explaining this phenomenon is to say that it is brought about by changing the weights each year in such a manner as to increase the importance of articles which have fallen in price and to

decrease the importance of those which have risen in price, and that this re-weighting causes trouble when the direction of movement is changed.

The manner of explanation is less important than the obvious fact that the chaining together of averages of link relatives of prices which have a tendency to return to normal gives an index series having a distinctly erroneous trend.

1926–1927; return of the price of every commodity to its original level.—In this period, we have a continuance of the tendency for high-priced commodities to fall and low-priced ones to rise. Just as in the preceding year, this results in a decline of the relative of aggregates and a rise in the average of relatives. The spliced index based upon the former is brought back to its starting point—unity—while the chained index based upon the averages of link relatives is now over 14 times as large as in the base year!

The desirability of studying index numbers by experimenting with actual data.—One is naturally suspicious of a method which yields an index series that fails to equal the base figure when all the individual items have returned to their respective values at the base date. Obviously, such an index number is entirely inappropriate for use with any type of data having a tendency to return to normal. It appears, however, that few if any investigations have been conducted to find how general this tendency actually is. Records of stock prices and commodity prices furnish an opportunity for testing the facts with two classes of data. We shall, then, follow the admirable example set by Professor Irving Fisher in his book entitled *The Making of Index Numbers* and attempt to ascertain experimentally the comparative behavior of index numbers computed by aid of the two different methods just mentioned. In order to avoid the danger of being misled by purely chance results, the same methods have been applied to several different samples. In this way, it is believed, one can ascertain whether, when

actual price quotations are used, indices based upon averages of link relatives tend to climb above those derived from aggregates, as it has been demonstrated that they do in the case of the hypothetical example presented in Table 20.

The stock-price records.—The first mass of data used in this experiment consists of records of the prices of industrial stocks at seven different dates. All but the last of these dates constitute important turning points in stock market averages. It is believed that, by selecting dates in this manner, the facts with which we are at present concerned can be demonstrated approximately as well as they could be by a study of records for every day or hour in the entire period covered.

For each stock, the following price-records were secured and entered on cards

High for November	1916
Low for December	1917
High for November	1919
Low for August	1921
High for March	1923
Low for October	1923
High for March	1925

Relative prices based upon the high for November, 1916 were also computed.

The 302 stocks covered include all those "industrial" shares important enough to be quoted in the price list given in Moody's *Rating Book* for 1917 and for which quotations could be obtained in or near each of the months specified. Stocks which, during the period, lost all their value were, for dates thereafter, rated at zero. The entire list of stocks is believed to give a very fair picture of the *universe* of reputable industrial shares in American markets.

The high for November, 1916 was taken as a starting point in every case. Allowance was made for all stock dividends and for every instance of splitting up of shares occurring after the date specified; hence the prices represent the

holdings at different dates of the owner of one share in November, 1916, provided he retained his stock throughout. No account was taken of ordinary cash dividends.

All price data for a given stock were entered on a single card. Relative prices were next computed for each stock, using the high for November, 1916 as a base.

The analysis was begun by classifying the cards thus:

1. Whether Common or Preferred.
2. On the basis of the highest market price per share recorded in November, 1916, all stocks being placed in one of the following three classes:
   (a) $0 to $39.99.
   (b) $40 to $104.99.
   (c) $105 or over.
3. On the basis of the nature of the business engaged in by the corporation, all companies being distributed among the following four classes:
   (a) Oil and Pipe Line Companies.
   (b) Coal Mining or Manufacturing Concerns.
   (c) Other Mining Companies.
   (d) Mercantile and Miscellaneous Corporations.

This method of classification would normally result in $2 \times 3 \times 4$ or 24 classes, but since, in this case, two of the classes contained no items, all the cards were in fact, distributed among 22 groups.

The selection of three samples.—The next step was to select three samples, each containing thirty of those cards upon which the price data for the respective stocks had been recorded. Since it was desired to make each sample represent the totality as closely as possible, the sampling was done on a proportional basis, the method used to secure proportionality being as follows:

Before selecting the sample cards from the various classes, the cards in each class of common stocks were arranged alphabetically. The first sample was made up of the cards on the tops, the second from the cards on the bottoms, and the third from the cards in the middles of the respective

piles. For each sample of common stocks, one card was selected out of every eleven cards in the class, but at least one card was taken from each class in which there were any items. When the cards in any class were exhausted, a card was chosen from the group most closely resembling the missing group in respect to kind of industry and value of stock.

Since there were approximately half as many preferred stocks as common stocks in the list, it was desired in selecting each sample, to choose 10 preferred stocks to go with the 20 common stocks. There were not enough preferred stocks in some of the classes to allow for adequate sampling; hence, all the cards were divided into 3 classes, according to their respective values in November, 1916. The classes were determined by the market value per share, and were as follows: under $40, $40 and under $105, and $105 and over. In making up the sample, one item was taken from the first class, 4 from the second, and 5 from the last class, making 10 items in all for each sample. The items were arranged alphabetically in each class, and the samples were selected by taking cards at approximately equal spacing throughout each pile.

The process just described gave us three samples, each selected in the same way, and each containing 30 stocks—the samples which are hereafter designated respectively as 1a, 1b, and 1c.

Computation of indices from fixed-base and from chain relatives.—The prices of all the various stocks in the three samples were converted to relatives, the base used for each stock being its price at the "High" of November, 1916. The relatives for the 30 stocks in each sample were then averaged arithmetically to obtain an average of relatives.

Next, prices relative to the preceding year were computed for each security; the link relatives thus obtained were averaged; and a spliced index series was computed from

these averages, using as a base the average of the "Highs" for November, 1916.

Comparative trends of fixed-base and chained relatives of stock prices.—The entries in Table 22 show the comparative trends of the index series obtained by the two methods. The comparison of the index obtained from the link relatives is made with the average of fixed-base relative prices rather than with a relative of aggregates because the average of fixed-base relatives, like the average of link relatives, gives

TABLE 22

A COMPARISON OF THE TRENDS OF INDICES OF STOCK PRICES COMPUTED FROM LINK RELATIVES WITH THE TRENDS OF FIXED BASE RELATIVES

Sample	Type of Relatives Used in Computation	High Nov. 1916	Low Dec. 1917	High Nov. 1919	Low Aug. 1921	High Mar. 1923	Low Oct. 1923	High Mar. 1925
1a	Fixed-base	1.00	0.667	1.995	0.632	1.011	0.775	1.290
	Link	1.00	0.667	1.124	0.645	1.097	0.750	1.286
1b	Fixed-base	1.00	0.843	1.746	1.009	2.018	1.881	3.559
	Link	1.00	0.843	1.784	1.070	1.779	1.364	1.895
1c	Fixed-base	1.00	0.655	1.131	0.634	0.981	0.789	1.191
	Link	1.00	0.655	1.182	0.673	1.084	0.799	1.253

equal weight to each stock in November, 1916. The aggregative method, on the other hand, gives to the shares that were high-priced in 1916 a weight so heavy as to make them dominate the behavior of the index numbers derived. For this reason, the course of the "link-relative index" series here presented, which weights all stocks equally, can not legitimately be compared with that of the relatives of aggregates.

A glance at Table 22 shows us that, in Sample 1a, the two

index numbers pursue about the same course; in Sample 1b, the fixed-base index runs far above the link-relative series; while, in Sample 1c, the fixed-base index holds a course slightly below the link-relative series. This experiment indicates that, in the case of stock prices, the trends of simple arithmetic averages of relatives having fixed bases are about the same as the trends of index numbers computed by chaining together averages of link relatives which have been derived by using as bases the respective preceding recorded dates. From the experiments made in Table 21, we may, then, infer that the reason for the similarity in the trends of the two types of index series shown in Table 22 is that stock prices do not tend either to go lower because they have declined or to recover after a decline; in other words, the behavior of stock prices resembles instead that of chance data. In this particular field, then, the use of the system of computing index numbers from link relatives seems to introduce no error not present when fixed-base relatives are used.

It may be worth while to explain at this point the striking difference in the behavior of the two indices derived from Sample 1b. The divergence is due almost wholly to the remarkable increase after 1921 in the price of Kresge stock. The spectacular rise of this stock affected the fixed-base index series more than it did the series derived from link relatives. In Sample 1b, therefore, the course of stock prices in general is represented better by the series of indices derived from link relatives than by the series depending upon fixed-base relatives.

The fact that both these series of indices of stock prices have very similar trends, is a matter of practical importance to makers of index numbers, for, because of the shifting characteristics of individual stocks, it is more convenient to derive index numbers by aid of link relatives than by most other methods of computation.

Commodity price data used in this analysis of index numbers.—In order to avoid basing conclusions upon one type of data only, it appears desirable to extend the laboratory study of the effect of various formulæ upon price indices to another field than that of stock prices—namely, to that of prices of commodities at wholesale.

The data utilized for this second study are the wholesale price quotations for 336 commodities given in the bulletins of the United States Bureau of Labor Statistics. With the exception of the base (the average price for the year 1919) the dates chosen are those which represent important turning points in the courses of the averages compiled by the Federal Bureau, namely, August, 1915, August, 1917, May, 1920, January, 1922, and August, 1925. The prices at all of these dates are compared with the average prices for 1919.

Most of the quotations used are those given in the original reports, but, in a number of instances, missing quotations have been estimated on the basis of changes in similar articles, or averages for the year have been used in lieu of the missing monthly figures. The supplying of occasional missing items made it possible to add substantially to the size of the totality dealt with, thus giving a better justification for our assumption that it approximates the universe of such data, and there is almost no possibility that the results have been materially affected by the presumably small errors in this limited number of estimates.

Since some of the commodities are quoted by the pound, some by the barrel, some by the dozen, and some by the yard, there appears to be no logic in computing averages of the original quotations. The prices for each commodity have, therefore, been reduced to relatives by dividing them by the average price for that commodity in the year 1919, and it is from the fixed-base relatives thus obtained that the sample items have been selected from which the index numbers have been computed.

Selection of samples from which to compute indices of commodity prices.—The prices for each commodity were entered on a single card; the 336 cards were then shuffled thoroughly; and three samples, each containing 31 items, were selected at random from the pack. These samples were labeled respectively 1a, 1b, and 1c.

Calculation of fixed-base and link relatives of commodity prices.—The simple arithmetic average of the 31 relatives for a given date was then calculated by adding the fixed-base relatives for the same date on all the cards in the sample and dividing the sum by the number of cards in the sample —namely, 31. In Table 23, the trends of the simple arithmetic averages of fixed-base relatives thus obtained are compared with the trends of averages of link relatives derived from the same samples.

The method of calculating the link-relative indices is as follows: Link relatives for each commodity were computed by using as a base the price in the period immediately preceding. For each date, the simple arithmetic average of the link relatives for the 31 commodities in a sample was then computed. For the three dates after 1919, the pieced-index series was computed by a series of multiplications. For the two dates preceding 1919, the computation was by successive divisions.

Comparative trends of indices of commodity prices derived from fixed-base relatives and from link relatives.—In Table 23, a comparison is made of the "chained indices" obtained by the method just described with averages of fixed-base relatives derived from the same samples. This comparison deals with three samples only—namely, 1a, 1b, and 1c, the results being so decisive that it was felt unnecessary to go to the work of extending the comparison to other samples.

An inspection of the figures in the table shows that, in all three instances, the indices derived from link relatives have trends rising more sharply than the trends of the aver-

ages of fixed-base relatives. Our analysis of the hypothetical figures in Table 21 showed that this condition would prevail wherever there is a tendency for prices to return to a normal. Evidently this tendency, though not found in stock prices, is present in the prices of commodities sold at wholesale; that is, the more a commodity has declined in price, the greater is the probability of a price recovery, and *vice versa.* The presence of this tendency in commodity prices evidently bars the use of the link-relative system in constructing index numbers for this field, for the results obtained by this method will tend to climb above the course of any truly representative index number.

TABLE 23

A COMPARISON OF THE TRENDS OF INDICES OF COMMODITY PRICES AT WHOLESALE DERIVED FROM LINK RELATIVES WITH AVERAGES OF FIXED-BASE RELATIVES

Sample	Type of Relatives Used in Computation	Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
1a	Fixed-base	0.477	0.925	1.000	1.231	0.786	0.867
	Link	0.421	0.800	1.000	1.231	0.803	0.931
1b	Fixed-base	0.508	0.959	1.000	1.246	0.719	0.839
	Link	0.431	0.866	1.000	1.246	0.785	0.955
1c	Fixed-base	0.476	0.762	1.000	1.253	0.805	0.820
	Link	0.424	0.750	1.000	1.253	0.966	1.009

SUMMARY

1. The "chained index numbers" discussed in this chapter, although computed by chaining together a series of link relatives having different bases, are intended to give a comparison with a fixed base just as do index numbers computed from relatives using a fixed base.

2. Link relatives are substituted for "fixed-base" relatives only because it is difficult or impossible to secure continuous, comparable data. They contain certain inherent tendencies to error against which the user must always be on guard.

3. Samples selected in the same way have equal probabilities of error; therefore the substitution of one sample for another does not necessarily increase the error in an index number. Such a substitution, however, often introduces into the series a jog which is entirely non-representative of the data.

4. The method of aggregates is the one at present most widely used in computing index numbers from changing samples. It is probably the best method for most purposes, but the data are not always in a form to permit of its convenient use. It is equivalent to averaging relatives after giving them a particular set of constant weights.

5. Index numbers derived from a series of link relatives climb above the true trend whenever the variables composing the data have a tendency to return to normal. Such indices tend to fall below the true trend whenever there is a tendency for a variable that has declined to fall further and for a variable that has risen to continue its rise. They have no bias in their trends when the variables are not influenced by their previous course but tend to fluctuate at random.

6. Prices of commodities at wholesale tend to return to normal; hence an index series derived from link relatives has an upward bias. Prices of stocks seem to fluctuate at random, having no tendency to return to normal; hence the use of index series derived from link relatives is permissible in this field.

QUESTIONS AND PROBLEMS

1. Is it usually possible to obtain price quotations covering a long period of years and applying to strictly comparable articles? How does your answer to this question bear on the question of constructing index numbers?

2. Do makers of index numbers of stock prices usually use the same sample of stocks over long periods of time? Why or why not?

3. Is it possible to construct an index number showing accurately the difference in the "cost of living" prevailing in

(a) The country and the city?
(b) Different nations?

4. In the case of many vegetables, price quotations are obtainable for but a few months in each year. How would you proceed if you wished to construct a monthly index of vegetable prices?

5. Does an index number derived by linking averages of price relatives tend to climb unduly when applied to prices which

(a) Fluctuate purely by chance?
(b) Tend to return to a rather well-defined normal?

6. In order to obtain representative samples, what devices can you suggest, other than the one actually used in the experiment described in this chapter, for selecting representative samples of

(a) Stocks?
(b) Commodities at wholesale?

7. A leading writer on index numbers asserts that the use of the arithmetic average imparts an upward bias to index numbers because, if a price increases from $80 to $100 there is a gain of 25 per cent, but when it again falls to $80, there is a loss of only 20 per cent. Does his argument apply to fixed-base relatives or to link relatives only?

8. Would the following procedure give a satisfactory index of the price of the most active stocks?

(a) Compute for each stock having heavy sales a price relative based upon the preceding period.
(b) Average the price relatives thus obtained.
(c) By multiplication, chain together the links thus formed.

Before answering this question, consider the usual relationship between activity and price movement as well as the results of chaining averages of relatives.

9. How could you construct an index of the movements of the most active stocks by employing the method of aggregates?

BIBLIOGRAPHICAL NOTE

Irving Fisher, in *The Making of Index Numbers* (1922), pp. 18–24, describes the construction of a chain-index series.

A brief but most enlightening article which deals with the averaging of link relatives is that by Frederick R. Macaulay in *The American Economic Review*, March, 1916, pp. 203–209.

The results of using link relatives instead of fixed-base relatives are discussed by Wesley C. Mitchell in *The Journal of Political Economy*, July, 1916, pp. 669–671, on pp. 81–93 of *Bulletin 284* of the United States Bureau of Labor Statistics (1921), and by Norman Crump in *The Journal of the Royal Statistical Society*, March, 1924, pp. 184–185.

CHAPTER VI

TYPES AND AVERAGES AS RELATED TO INDEX NUMBERS

Need for understanding types and averages.—As pointed out in earlier chapters, the only true test of the merit of a formula for an index number is the closeness with which the series of indices derived by its application to sample data follows the course of that series of relatives, computed from the totality of items, which exactly answers the question asked. Under ordinary circumstances, however, the maker of index numbers has no chance to apply this test, for complete data are not obtainable. The best that he can do is to learn as much as possible about the behavior of different types of relatives and index numbers and, by deductive reasoning and experimentation on a limited scale, determine which formula, when applied to the sample data, will produce the series of index numbers best adapted to the particular end in view.

Now, in the process of using any sample data for the purpose of constructing an index number, it is always necessary either to select some typical item or to compute some kind of an average of the individual items in the sample.

Types and averages used in relative comparison.—Thus far it has been tacitly assumed that the arithmetic is the only kind of average used in the construction of index numbers. As a matter of fact, it is the one used in the large majority of cases, but there are outstanding authorities who believe that, for certain purposes, the arithmetic average

is distinctly inferior. It therefore behooves the student of index numbers to familiarize himself with some of the more common types and averages occasionally used in their construction.

Index numbers may be constructed by aid of an infinite variety of formulæ. Most of the processes, however, make use of combinations of a limited number of types or averages. The simplest of these types is the *median.* When all the items under consideration are arrayed in order of size, it is the middle item of the series. Suppose, for example, that we are required to find the median of 9, 3, 8, 2, 6, 11, 2, 3, and 7. The first step is to array the items thus: 2, 2, 3, 3, 6, 7, 8, 9, 11. The middle item (namely, 6) is the *median.*

In some instances, the various items are deemed to differ in importance and hence it is desired to assign to each a fitting weight. If, for instance, in the above example, the first three items in the list were all considered three times as important as any of the others, the process of weighting would be accomplished by repeating each of these items three times. The original series would then become 9, 9, 9, 3, 3, 3, 8, 8, 8, 2, 6, 11, 2, 3, and 7, and the array would appear as follows: 2, 2, 3, 3, 3, 3, 6, 7, 8, 8, 8, 9, 9, 9, 11. The *weighted median* is the middle item of this new array —namely, 7.

One peculiarity of the median which makes it very useful for certain purposes is that it is unaffected by changes in the size of the extreme items. In the given example, if the two smallest items were each reduced to 1, or increased to 3 or 5, and if the largest items were decreased to 7 or increased to an unlimited extent, the median item would still be 6. This characteristic is not possessed by *averages* as distinguished from *types.*

When the number of items is small, the central items may be considerably scattered, causing the median to behave in

an erratic manner. A procedure frequently adopted to overcome this tendency is to use as a type not the median but an arithmetic average of a small group of items in the center of the array. The difficulty and the remedy may both be illustrated in connection with the example of the simple or unweighted median cited above. In the instance given the median is 6. Let us suppose that, in the next year, two new items, namely 1 and 3, are added. Then the array becomes 1, 2, 2, 3, 3, 3, 6, 7, 8, 9, 11. The median is now the sixth item, or 3. The addition of the two small items has, then, resulted in halving the median—a most radical change. If, instead, we use as a type in each case the arithmetic average of the three middle items, the violence of the change is much lessened. In the first instance, the average of 3, 6, and 7 is $5\frac{1}{3}$. In the second case, the average of 3, 3, and 6 is 4. The decline in the averages of the three middle items is, then, only 25 per cent as compared to the 50 per cent diminution recorded when the median items were taken as types.

Index numbers are generally computed by aid of one of the following three kinds of averages, the *arithmetic,* the *harmonic* or the *geometric.* The one most commonly used is the *arithmetic* average or mean—the average which, as everyone knows, is computed by adding together all the items and dividing the sum by their number. For example, the arithmetic mean of 3, 4, and 6 is $\frac{3+4+6}{3}$ or 4.33.

The *harmonic average* is sometimes referred to as the arithmetic average reversed. To find the harmonic average of 3, 4, and 6, the reciprocals of the numbers are first obtained, that is, the numbers are made denominators of fractions having 1 as the numerator, and the fractions are added, thus:

$$\tfrac{1}{3}+\tfrac{1}{4}+\tfrac{1}{6}=$$

$$\tfrac{4}{12}+\tfrac{3}{12}+\tfrac{2}{12}=\tfrac{9}{12}=\tfrac{3}{4}$$

The sum of the fractions is now divided by the number of items, namely 3, which gives $\frac{3}{4} \div 3 = \frac{1}{4}$. The reciprocal of the quotient, or, in other words, the quotient inverted, is the harmonic average. In this instance: $1 \div \frac{1}{4} = 4$ which is the harmonic average of 3, 4, and 6. This result may be contrasted with the arithmetic average which is 4.33. In practice, the computation can usually be facilitated by using decimals instead of common fractions. Thus:

Reciprocal of 3 = 0.333
Reciprocal of 4 = 0.250
Reciprocal of 6 = 0.167

$3\overline{)0.750}$

0.250

Reciprocal of 0.250 = 4.00

The *geometric average* is found by multiplying all the items together and extracting the root corresponding to the number of items. Thus, with two items, the square root is found; when three items are used the cube root is required; and with four items it is necessary to extract the fourth root.[1]

The *geometric average* of 3, 4, and 6 is arrived at thus:

$$3 \times 4 \times 6 = 72$$

$$\sqrt[3]{72} = 4.16$$

When there are more than two items, the use of logarithms greatly simplifies the work of computation. For example, the easy way to find the geometric average of 248, 27, 896, and 1,172 is as follows:

Logarithm of 248 = 2.3945
Logarithm of 27 = 1.4314
Logarithm of 896 = 2.9523
Logarithm of 1,172 = 3.0689

$4\overline{)9.8471}$

2.4618

[1] When it is desired to give different weights to the items entering into the geometric average, this is accomplished by repeating the items. Thus if, in the present example, it is desired to give to the first item a weight of four and to the second item a weight of two, the weighted geometric average would be $\sqrt[7]{3 \times 3 \times 3 \times 3 \times 4 \times 4 \times 6}$.

2.4618 is the logarithm of 289.6 which is the geometric average of the four numbers.

One of the weaknesses of the geometric average is the fact that it can be used only in those cases in which all the items to be averaged are positive, the presence of a single negative item destroying its significance.

Relative sizes of the arithmetic, geometric, and harmonic averages.—By averaging 3, 4, and 6 in the three ways mentioned, we obtained the following results:

Arithmetic average = 4.33
Geometric average = 4.16
Harmonic average = 4.00

It will be observed that the *geometric* average is *smaller* than the *arithmetic* average but *larger* than the *harmonic* average. Experiment will show that this relationship invariably holds except when all the items are the same size, in which case all the averages are identical.

The greater the scatter in the sizes of the various items, the greater will be the divergence of the other two averages from the geometric average.

The fact that the geometric average always lies between the arithmetic and the harmonic averages explains the peculiar changes which appear in the relationships of averages of relatives whenever the base used in computing the relatives is shifted. This change in the position of the three averages is illustrated by Table 24—a table which deals with the prices of certain hypothetical stocks.

It is clear that changing the base used in computing the relatives has a striking effect on both the arithmetic and harmonic averages of relatives. With the base located in 1915, the arithmetic average shows a sharply *rising* trend; while, with the base in 1925, it shows, instead, a trend *falling* just as sharply. The reason for this difference is that the use of the prices of 1915 as bases multiplies by

16 the importance of the rise in American Textile as compared to the fall in United Sugar.

TABLE 24

COMPARATIVE CHARACTERISTICS OF DIFFERENT AVERAGES OF RELATIVE STOCK PRICES

	PRICE PER SHARE		
	1915	1920	1925
PRICES PER SHARE:			
United Sugar	$128	$64	$32
Central Lead	32	32	32
American Textile	8	16	32
RELATIVE PRICES:	Relatives: Base 1915 = 100		
United Sugar	100	50	25
Central Lead	100	100	100
American Textile	100	200	400
AVERAGES OF RELATIVES:			
Arithmetic	100	117	175
Geometric	100	100	100
Harmonic	100	86	57
RELATIVE PRICES:	Relatives: Base 1925 = 100		
United Sugar	400	200	100
Central Lead	100	100	100
American Textile	25	50	100
AVERAGES OF RELATIVES:			
Arithmetic	175	117	100
Geometric	100	100	100
Harmonic	57	86	100

Geometric average not affected by changing the base for the individual relatives.—The harmonic average, being the antithesis of the arithmetic average, behaves, as might be expected, in exactly the reverse manner. The geometric average, however, is not affected in the least by the change

of base, but shows a horizontal trend in each case, for it gives the same weight to the doubling of the price of American Textile as to the halving of the price of United Sugar; in other words, it gives equal weight to equal *proportional* changes. In this respect, it is unlike the arithmetic average which gives equal weight to equal *absolute* changes in the items averaged.

The geometric average is seldom useful for obtaining relative comparisons from *complete* data, for the formula does not correspond to any common type of question.[2] However, many statisticians have considered it an average most useful in the computation of index numbers. The fact that changing the base used for the individual relatives does not change the index number obtained from a geometric average of these relatives has seemed to them to indicate that index numbers computed by aid of the geometric average have especial merit. Such index numbers conform to Professor Pierson's "time-reversal test"—a test which requires that shifting from one point to another the base used in the calculation of the relatives must not change the ratio between the index numbers derived by averaging the relatives at these respective points.

As a matter of fact, the belief that conformity to the time-reversal test is an indication of "merit" has no justification in fact. This characteristic of the geometric average is, nevertheless, something worth keeping in mind, for it aids one in forecasting the probable behavior of index numbers derived by use of this average, and hence helps one to judge whether or not such an index series is the one likely to fit most closely the course of that series of relatives exactly answering the question in mind—relatives neces-

[2] The reader can readily convince himself of the truth of this assertion by first writing down some questions of the type for which he would expect index series to give answers, and then framing in algebraic terms the formulae which, if applied to complete data, would yield the correct indices.

sarily calculated from complete data for this particular purpose.

Explanation of the supposed "bias" of certain index numbers.—That the trend of geometric averages of relatives is not affected by changing the base used for the individual relatives is not the only interesting fact revealed by the figures in Table 24, for a comparison of the three averages recorded in the table shows that, except in the base year, the arithmetic average is always larger than the geometric average and the harmonic average is always smaller than the geometric average. This peculiar relationship of these three averages has already been referred to as one which universally prevails except when all the items averaged are of the same size. Now all relatives having a common base must obviously be of the same size at this base date, hence, if the three averages mentioned are computed from the same set of relatives, it will almost inevitably happen that, as Professor Irving Fisher has pointed out,[3] the curves representing the three types of index numbers will diverge from the base date like the prongs of a three-tined fork—the arithmetic average being above, the geometric in the middle, and the harmonic below.

The tines always point away from the base; hence, if the base is the first date, they point forward in time; if the base is at the end of the period they point backward in time; and, if the base is in the middle of the series, there are two forks pointing in opposite directions from the base which forms a focal point. It is this phenomenon which has given rise to the assertion stressed so strongly by Professor Fisher that the arithmetic average has an "upward bias" and the harmonic average has a "downward bias." [4]

As to the respective positions occupied by these three averages, there can be no doubt. One may, however, ques-

[3] *The Making of Index Numbers*, p. 92.
[4] *Ibid.*, p. 86.

tion seriously the advisability of applying to either of the two averages the term "bias"—a term which is commonly used in a derogatory sense—for its use implies that the central position is in some way associated with accuracy or merit. This hypothesis needs statistical verification before being accepted. Without such verification, we cannot be sure that the best fit to the course of the required series of relatives computed from complete data may not be given by that average of the sample forming the lowest tine of the fork. Perhaps, on the other hand, the upper tine will give the smallest error. To guess about the matter is hazardous.

Arithmetic, geometric, and harmonic averages of original data compared.—It is worth while also to observe that the three-tined fork arrangement of the arithmetic, the geometric, and the harmonic averages characterizes only one particular kind of index numbers, namely that variety derived by averaging *individual relatives having a common base,* and it is a mistake to assume that this is the only type of index numbers worthy of consideration. Averages of the original items—measured perhaps in terms of bushels, feet, or dollars—are frequently much easier to compute, and there is no reason for sentencing them to a position of inferiority until they have been tried. Before proceeding further, therefore, let us see how the three kinds of averages behave when applied directly to the same prices of the same stocks recorded in Table 24—their application in the present instance, however, being to the dollar values and not to the relatives computed from these values.

The original dollar values have been averaged according to the three different formulæ and each of the averages thus obtained has been translated into terms of dollars on the base 1915. Table 25 presents the results of these operations.

It is now made obvious that, when the averaging is done first and the relatives are computed later, the "inherent

tendency" for the arithmetic average to have an upward "bias" is not so apparent; indeed, in the example here given, we find that it is the series of *harmonic* averages which shows the sharp upward trend, while the relatives representing the *arithmetic* average are the ones that show the heavy decline.

TABLE 25

RELATIVES DERIVED FROM DIFFERENT AVERAGES OF THE ORIGINAL DATA

	PRICE PER SHARE:		
	1915	1920	1925
STOCK:			
United Sugar	$128	$64	$32
Central Lead	32	32	32
American Textile	8	16	32
Total	$168	$112	$96
AVERAGE OF PRICES PER SHARE:			
Harmonic	18	27	32
Geometric	32	32	32
Arithmetic	56	36	32
RELATIVES DERIVED FROM AVERAGE PRICES:	Base: 1915 = 100		
Harmonic	100	150	178
Geometric	100	100	100
Arithmetic	100	66	57

Converting items to relatives is equivalent to giving new weights to the items.—Furthermore, the results recorded in Table 25 contrast sharply with those entered in the middle of Table 24—even though both sets of relatives are on the base 1915 and are derived from identical data by use of the same formula—for, in Table 25, the relatives derived from the arithmetic averages of the original data show a rapidly declining trend, while the arithmetic averages of the rela-

tives recorded in Table 24 show just as sharp a rise. The reason for this striking difference is that the method used in Table 25 gives to American Textile, which is a declining stock, little weight because it is low priced and to United Sugar much weight because it is highly valued. The method used in the first part of Table 24, however, gives equal weight to both stocks in 1915. The weighting has the effect of multiplying the prices per share of American Textile by 16 before averaging them with the quotation for United Sugar. Again we have an illustration of the fact that converting numbers into relatives is equivalent to giving them new weights.

If we compare the relatives of harmonic averages in the last section of Table 25 with the harmonic averages of relatives appearing in the first part of Table 24, we find that, even though both series are on the base 1915, the trends are exactly opposite to each other. Furthermore the trends are, in both cases, the inverse of those of the series computed by the rise of arithmetic averages. This last phenomenon is due to the fact that the harmonic averages of the relatives get a weighting just the reverse of that applying to the arithmetic averages of the relatives.

The arrangement brought out in Table 24, namely that of the three-tined fork with the arithmetic average at the top and the harmonic at the bottom, is, then, clearly the result of the particular method used—that of reducing the original prices to relatives on a common base and averaging these relatives. When the original data themselves are averaged, as in Table 25, the comparative slopes of the series of relatives derived by applying the three kinds of averages cannot be predicted so easily. Evidently, then, the slope of a series of index numbers is largely dependent upon the nature of the averages used in the computation of the indices.

The trends of five different kinds of index numbers computed from actual prices of stocks.—The discussion thus far

in this chapter has been confined entirely to hypothetical data. Let us now again turn to the experimental method of attack and compare the results obtained by the application of various types and averages to the prices of the same three samples of industrial stocks, the selection of which was described in the last chapter. The five different index numbers, all on the same base—namely the November, 1916, high—are presented in Table 26.

TABLE 26

A COMPARISON OF INDICES COMPUTED BY DIFFERENT AVERAGING PROCESSES FROM THE PRICES OF THREE SAMPLES CONTAINING THIRTY STOCKS EACH

SAMPLE	FORMULA USED	INDICES OF STOCK PRICES AT THE FOLLOWING DATES						
		Nov. 1916 high	Dec. 1917 low	Nov. 1919 high	Aug. 1921 low	Mar. 1923 high	Oct. 1923 low	Mar. 1925 high
1a	Means of relatives	100.00	66.70	119.46	63.19	101.13	77.52	128.99
	Relatives of aggregates	100.00	72.93	104.92	61.50	92.96	74.11	105.70
	Medians of prices per share	100.00	63.73	102.91	36.75	72.70	49.60	77.39
	Medians of relatives	100.00	64.60	93.80	60.65	87.50	74.65	84.80
	Geometric averages of prices per share	100.00	61.36	97.26	50.42	78.79	50.84	76.46
1b	Means of relatives	100.00	84.34	174.63	100.87	201.84	188.06	335.95
	Relatives of aggregates	100.00	73.22	115.49	69.97	113.68	97.86	135.89
	Medians of prices per share	100.00	74.42	109.66	63.13	99.80	87.62	101.97
	Medians of relatives	100.00	72.55	98.85	71.60	87.35	81.85	90.90
	Geometric averages of prices per share	100.00	69.36	123.79	64.72	102.59	63.67	90.05
1c	Means of relatives	100.00	65.51	113.08	63.43	98.15	78.93	119.10
	Relatives of aggregates	100.00	68.05	110.40	64.36	99.69	85.83	95.49
	Medians of prices per share	100.00	66.90	115.24	56.04	99.74	78.02	119.21
	Medians of relatives	100.00	66.90	91.50	66.80	89.90	77.85	95.80
	Geometric averages of prices per share	100.00	59.92	96.36	49.37	75.12	49.41	68.88

It may be well at this point to explain one omission. It has not been deemed worth while to compare the geometric average of the relatives with the arithmetic average of the relatives, inasmuch as we know in advance that, for inherent

mathematical reasons, the arithmetic average of the relatives must, for every date except the base, November, 1916, be higher than the geometric average of the relatives.

One can scarcely help being impressed with the great disparity between the indices for the same date derived by a common formula from the different samples. Evidently a sample of far more than thirty items is essential if chance variations are to be eliminated and statistical regularity is to be attained. However, despite the diversity in results obtained from different samples, there is, nevertheless, comparative regularity in the positions of the different types of indices when those derived from each of the three samples are arrayed according to their respective trends. There appears, indeed, to be a definite tendency in each sample to form a four- or five-tined fork with the tines diverging from November, 1916. The order of the tines is shown by the sequence of formulæ in the entries in Table 26. The arithmetic means of the relatives tend to rise above all the others, having a rather sharp upward trend. Two out of three of the relatives of aggregates show a moderate upward trend. The medians of actual prices indicate a trend not far from horizontal. The medians of relatives all have a moderate decline. The geometric averages of actual prices indicate a rather marked downward trend.

Obviously, the fact that index numbers computed by different formulæ show such radically different trends proves that the question of which formula to use in computing an index number is not one of minor importance, for, by substituting one formula for another, we can secure index numbers having almost any type of trend desired.

Why do different formulæ, when applied to a given sample of data, produce such radically different index numbers? Do the differences in trend result from qualities inherent in the formulæ used in averaging or do they instead merely reflect peculiarities of this particular kind of data? To throw

light on this question, Table 27 and 28 have been constructed.

Computation of aggregates and relatives of stock prices. —The first step in classifying the data for tabulation was to separate the 302 stocks for which quotations were secured into two main groups—common and preferred. Each of these groups was next divided into three categories, the assignment depending on the highest value per share attained by the stock in November, 1916. Each of the classes thus obtained was again subdivided into four subclasses on the basis of the nature of the business engaged in by the corporation. The totals in Table 27 were arrived at by adding together the individual prices of one share of each stock in the given class. The figures in Table 28 were secured by reducing the totals in Table 27 to relatives on the base November, 1916.

Diversity in the movements of different classes of stocks. —A study of these relatives reveals some interesting facts concerning the nature of this particular group of data. We see, for instance, that, when the computation is made in the manner stated, the total of all common stocks increased 21 per cent, while the total for all preferred stocks declined 9 per cent during the 8½ year period. The common stocks of "Other Mining Companies" suffered a loss of 57 per cent as contrasted with a gain of 125 per cent for "Mercantile and Miscellaneous" shares. The common stocks of the "Coal Mining and Factories" group show an increase of 11 per cent while the preferred stocks of the same kind register a 12 per cent decline.

Despite a heavy loss in the value of the mining stocks in the group, the total of those common shares selling below $40 in November, 1916, shows the startling increase of 141 per cent, the low-priced "Mercantile and Miscellaneous" common stocks actually increasing by 700 per cent. By contrast, the common shares selling above $105 at the date

TABLE 27

THE TOTAL VALUES IN DOLLARS OF INDUSTRIAL STOCKS OF DIFFERENT CLASSES CALCULATED ON THE BASIS OF ONE SHARE OF EACH STOCK IN NOVEMBER, 1916

Price per Share in Nov., 1916	Industry	Number of Stocks	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High
			COMMON AND PREFERRED STOCKS						
All values	**All industries**	**302**	**30,365**	**21,744**	**32,750**	**20,045**	**31,637**	**25,588**	**33,752**
	Oil and pipe lines	**28**	**6,850**	**5,399**	**8,401**	**5,414**	**9,863**	**7,706**	**9,964**
	Coal and factories	**177**	**16,758**	**11,873**	**18,409**	**11,092**	**15,876**	**13,281**	**16,860**
	Other mining	**62**	**3,754**	**2,115**	**2,514**	**1,388**	**2,383**	**1,479**	**1,857**
	Mercantile and miscellaneous	**35**	**3,003**	**2,357**	**3,426**	**2,151**	**3,515**	**3,122**	**5,071**
$0 to $39.99	**All industries**	**67**	**1,214**	**870**	**2,321**	**1,067**	**1,998**	**1,545**	**3,021**
	Oil and pipe lines	5	103	87	440	137	217	128	229
	Coal and factories	22	493	373	1,103	474	767	562	1,100
	Other mining	31	527	233	318	148	298	178	233
	Mercantile and miscellaneous	9	191	177	460	308	716	677	1,459
$40.00 to $104.99	**All industries**	**126**	**9,377**	**6,490**	**11,364**	**6,307**	**9,655**	**7,460**	**9,741**
	Oil and pipe lines	6	432	302	782	360	655	475	622
	Coal and factories	90	6,876	4,830	8,622	4,929	7,311	5,765	7,371
	Other mining	19	1,316	665	933	506	905	596	775
	Mercantile and miscellaneous	11	753	693	1,027	512	784	624	973
$105.00 or over	**All industries**	**109**	**19,674**	**14,384**	**19,065**	**12,671**	**19,984**	**16,583**	**20,990**
	Oil and pipe lines	17	6,315	5,010	7,179	4,917	8,991	7,103	9,113
	Coal and factories	65	9,389	6,670	8,684	5,689	7,798	6,954	8,389
	Other mining	12	1,911	1,217	1,263	734	1,180	705	849
	Mercantile and miscellaneous	15	2,059	1,487	1,939	1,331	2,015	1,821	2,639
			COMMON STOCKS						
All values	**All industries**	**204**	**20,677**	**13,900**	**23,348**	**13,168**	**22,931**	**17,949**	**24,966**
	Oil and pipe lines	**25**	**6,583**	**5,195**	**7,977**	**5,176**	**9,560**	**7,466**	**9,645**
	Coal and factories	**103**	**9,207**	**5,755**	**11,144**	**5,719**	**9,145**	**7,389**	**10,192**
	Other mining	**56**	**3,225**	**1,715**	**2,703**	**1,061**	**1,930**	**1,094**	**1,396**
	Mercantile and miscellaneous	**20**	**1,662**	**1,235**	**2,154**	**1,212**	**2,296**	**2,022**	**3,733**

$0 to $39.99	**All industries**	**62**	**1,195**	**801**	**2,178**	**976**	**1,873**	**1,456**	**2,878**
	Oil and pipe lines	5	103	87	440	137	217	128	229
	Coal and factories	18	385	311	968	395	659	485	976
	Other mining	31	527	233	318	148	298	178	233
	Mercantile and miscellaneous	8	180	170	452	296	699	665	1,440
$40.00 to $104.99	**All industries**	**85**	**5,980**	**3,797**	**7,686**	**3,860**	**6,437**	**4,822**	**6,576**
	Oil and pipe lines	4	303	220	466	212	462	323	413
	Coal and factories	60	4,193	2,663	5,813	3,006	4,837	3,729	5,006
	Other mining	16	1,127	554	794	425	746	492	639
	Mercantile and miscellaneous	5	357	360	613	217	392	278	518
$105.00 or over	**All industries**	**57**	**13,502**	**9,302**	**13,484**	**8,332**	**14,621**	**11,671**	**15,512**
	Oil and pipe lines	16	6,177	4,888	7,071	4,827	8,881	6,993	9,003
	Coal and factories	25	4,629	2,781	4,363	2,318	3,649	3,175	4,210
	Other mining	9	1,571	928	961	488	886	424	524
	Mercantile and miscellaneous	7	1,125	705	1,089	699	1,205	1,079	1,775

PREFERRED STOCKS

All values	**All industries**	**98**	**9,688**	**7,844**	**9,402**	**6,877**	**8,706**	**7,639**	**8,786**
	Oil and pipe lines	**3**	**267**	**204**	**424**	**238**	**303**	**262**	**319**
	Coal and factories	**74**	**7,551**	**6,118**	**7,265**	**5,373**	**6,731**	**5,892**	**6,668**
	Other mining	**6**	**529**	**400**	**441**	**327**	**453**	**385**	**461**
	Mercantile and miscellaneous	**15**	**1,341**	**1,122**	**1,272**	**939**	**1,219**	**1,100**	**1,338**
$0 to $39.99	**All industries**	**5**	**119**	**69**	**143**	**91**	**125**	**89**	**143**
	Oil and pipe lines	0	0	0	0	0	0	0	0
	Coal and factories	4	108	62	135	79	108	77	124
	Other mining	0	0	0	0	0	0	0	0
	Mercantile and miscellaneous	1	11	7	8	12	17	12	19
$40.00 to $104.99	**All industries**	**41**	**3,397**	**2,693**	**3,678**	**2,447**	**3,218**	**2,638**	**3,165**
	Oil and pipe lines	2	129	82	316	148	193	152	209
	Coal and factories	30	2,683	2,167	2,809	1,923	2,474	2,036	2,365
	Other mining	3	189	111	139	81	159	104	136
	Mercantile and miscellaneous	6	396	333	414	295	392	346	455
$105.00 or over	**All industries**	**52**	**6,172**	**5,082**	**5,581**	**4,339**	**5,363**	**4,912**	**5,478**
	Oil and pipe lines	1	138	122	108	90	110	110	110
	Coal and factories	40	4,760	3,889	4,321	3,371	4,149	3,779	4,179
	Other mining	3	340	289	302	246	294	281	325
	Mercantile and miscellaneous	8	934	782	850	632	810	742	864

TABLE 28

RELATIVE CHANGES IN THE TOTAL VALUES OF INDUSTRIAL STOCKS OF DIFFERENT CLASSES CALCULATED ON THE BASIS OF ONE SHARE OF EACH STOCK AT THE BASE DATE, NOVEMBER, 1916

Price per Share in Nov. 1916	Industry	Number of Stocks	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High
			COMMON AND PREFERRED STOCKS						
All values	**All industries**	**302**	**100**	**72**	**108**	**66**	**104**	**84**	**111**
	Oil and pipe lines	**28**	**100**	**79**	**123**	**79**	**144**	**112**	**145**
	Coal and factories	**177**	**100**	**71**	**110**	**66**	**95**	**79**	**101**
	Other mining	**62**	**100**	**56**	**67**	**37**	**63**	**39**	**49**
	Mercantile and miscellaneous	**35**	**100**	**78**	**114**	**72**	**117**	**104**	**169**
$0 to $39.99	**All industries**	**67**	**100**	**66**	**177**	**81**	**152**	**118**	**230**
	Oil and pipe lines	5	100	84	427	133	211	124	222
	Coal and factories	22	100	76	224	96	156	114	223
	Other mining	31	100	44	60	28	57	34	44
	Mercantile and miscellaneous	9	100	93	241	161	375	354	764
$40.00 to $104.99	**All industries**	**126**	**100**	**96**	**121**	**67**	**103**	**80**	**104**
	Oil and pipe lines	6	100	78	181	83	152	110	144
	Coal and factories	90	100	70	125	72	106	84	107
	Other mining	19	100	51	71	38	69	45	59
	Mercantile and miscellaneous	11	100	92	136	68	104	83	129
$105.00 or over	**All industries**	**109**	**100**	**73**	**97**	**64**	**102**	**84**	**107**
	Oil and pipe lines	17	100	79	114	78	142	112	144
	Coal and factories	65	100	71	92	61	83	74	89
	Other mining	12	100	64	66	38	62	37	44
	Mercantile and miscellaneous	15	100	72	94	65	98	88	128
			COMMON STOCKS						
All values	**All industries**	**204**	**100**	**67**	**113**	**64**	**111**	**87**	**121**
	Oil and pipe lines	**25**	**100**	**79**	**121**	**79**	**145**	**113**	**146**
	Coal and factories	**103**	**100**	**63**	**121**	**62**	**99**	**80**	**111**
	Other mining	**56**	**100**	**53**	**64**	**33**	**60**	**34**	**43**
	Mercantile and miscellaneous	**20**	**100**	**74**	**130**	**73**	**138**	**122**	**225**

$0 to $39.99	**All industries**	**62**	**100**	**67**	**182**	**82**	**157**	**122**	**241**
	Oil and pipe lines	5	100	84	427	133	211	124	222
	Coal and factories	18	100	81	251	103	171	126	254
	Other mining	31	100	44	60	28	57	34	44
	Mercantile and miscellaneous	8	100	94	251	164	388	369	800
$40.00 to $104.99	**All industries**	**85**	**100**	**63**	**129**	**65**	**108**	**81**	**110**
	Oil and pipe lines	4	100	73	154	70	152	107	136
	Coal and factories	60	100	64	139	72	115	89	119
	Other mining	16	100	49	70	38	66	44	57
	Mercantile and miscellaneous	5	100	101	172	61	110	78	145
$105.00 or over	**All industries**	**57**	**100**	**69**	**100**	**62**	**108**	**86**	**115**
	Oil and pipe lines	16	100	79	114	78	144	113	146
	Coal and factories	25	100	60	94	50	79	69	91
	Other mining	9	100	59	61	31	56	27	33
	Mercantile and miscellaneous	7	100	63	97	62	107	96	158

PREFERRED STOCKS

All values	**All industries**	**98**	**100**	**81**	**97**	**71**	**90**	**79**	**91**
	Oil and pipe lines	**3**	**100**	**76**	**159**	**89**	**113**	**98**	**119**
	Coal and factories	**74**	**100**	**81**	**96**	**71**	**89**	**78**	**88**
	Other mining	**6**	**100**	**76**	**83**	**62**	**86**	**73**	**87**
	Mercantile and miscellaneous	**15**	**100**	**84**	**95**	**70**	**91**	**82**	**100**
$0 to $39.99	**All industries**	**5**	**100**	**58**	**120**	**76**	**105**	**75**	**120**
	Oil and pipe lines	0	100	0	0	0	0	0	0
	Coal and factories	4	100	57	125	73	100	71	115
	Other mining	0	100	0	0	0	0	0	0
	Mercantile and miscellaneous	1	100	64	73	109	154	109	173
$40.00 to $104.99	**All industries**	**41**	**100**	**79**	**108**	**72**	**95**	**78**	**93**
	Oil and pipe lines	2	100	64	245	115	150	118	162
	Coal and factories	30	100	81	105	72	92	76	88
	Other mining	3	100	59	74	43	84	55	72
	Mercantile and miscellaneous	6	100	84	104	74	99	87	115
$105.00 or over	**All industries**	**52**	**100**	**82**	**90**	**70**	**87**	**80**	**89**
	Oil and pipe lines	1	100	88	78	65	80	80	80
	Coal and factories	40	100	82	91	71	87	79	88
	Other mining	3	100	85	89	72	86	83	96
	Mercantile and miscellaneous	8	100	84	91	68	87	79	92

mentioned made an average gain of only 15 per cent. Preferred stocks of the lowest priced group show a 20 per cent rise while the highest priced class were 11 per cent lower in March, 1925, than in November, 1916.

It appears, then, that, in November, 1916, that part of the public interested in the stock market underestimated the income that the low-priced stocks were going to yield to their owners in the near future; hence these stocks rose in value during the nine years covered by the study. This undervaluation occurred in the case of the lower grade stocks in three of the four classes into which the list of companies has been divided. In view of the rather haphazard relationship existing between investment and total par value of stock, it is not easy to offer any very logical explanation of the situation. The facts, nevertheless, seem to be clear. It is not, however, safe to assume that the lower priced shares are always the better investments, for it has not yet been proved that undervaluation of such shares has characterized all or most other periods.

The wide variations occurring in the behavior of the prices of the different classes of the 302 stocks included in the study, and not the nature of the formulæ used in the computation, appear to furnish the real explanation of the very considerable degree of dispersion prevailing among the various index numbers recorded in Table 26. A study of the material in some detail will make the truth of this statement more evident.

Reasons for variance in behavior of stock-price index numbers computed by different formulæ.—It has been shown on a previous page that, as compared to an index number derived from aggregates, an index number consisting of the arithmetic mean of relatives gives heavier weight to the stocks selling at a low price in the base year. If, then, the stocks having low prices at that date rise in price more rapidly than do the high-priced stocks, the means of relatives

will, when plotted, lie above the means of aggregates. Since, as we have seen, the stocks selling under $40 per share more than doubled in value during the period, while, in general, the higher priced stocks rose but slightly, the reason why the index series consisting of the means of relatives occupies the upper level is, then, obvious. The relative positions of these two indices truthfully indicates the fact that the investor who, in November, 1916, put equal amounts of money into the various kinds of shares fared better than did the man who bought an equal number of shares of each stock.

Both medians and geometric averages are less affected by stocks having spectacular ascents in value than are arithmetic averages, hence, when we find the stocks in the mercantile and miscellaneous group rising in 1925 to seven times their 1916 value, we must naturally expect both index numbers dependent upon arithmetic averages to rise above the other indices—an expectation which is fulfilled.

The reasons for the other relationships are less apparent. Why, for example, do the medians of the actual prices per share tend to run higher than the medians of the relatives? Probably the relationship between the trends of these two types of medians is purely accidental, being due to peculiarities of the particular data used and the specific period covered.

Another puzzling question is why the series of geometric averages of actual stock prices shows a downward slope. Our experiment with hypothetical data proved that there was nothing inherent in the geometric formula to make the indices behave thus. The information is scarcely sufficient to enable us to answer the question in detail. However, since, as we previously pointed out, the ratio of the arithmetic average to the geometric average increases as the dispersion of the items grows larger, it is evident that the dispersion of the stock prices increased during the period covered, thus causing the geometric averages to show a down-

ward trend as compared to the arithmetic averages. This relationship might be reversed if the prices of other stocks or the quotations on the same stocks at another period were used in the computation. Presumably, the slopes of all five of the index numbers are determined wholly by the nature of the data used and not by any qualities inherent in the respective formulæ employed in calculating the indices.

The figures just presented serve to make clear the different effects produced by applying various types or averages to the prices of stocks. Do the same principles established for stock prices hold for commodity prices?

Especial interest is attached to this question because Professor Irving Fisher has devoted most of his book on *The Making of Index Numbers* to a study of index numbers of commodity prices at wholesale. In this study, he lays especial stress on the superiority of the weighted over the unweighted arithmetic average. For example, on page 87, he says: "The arithmetic average [of relatives] exhibits an *inherent tendency* to exaggeration, a 'bias', such that, in the instance cited, it yields a result probably too high by about 5½ per cent." He states further: "This inherent tendency in the arithmetic type always exists irrespective of the method of weighting used." Later however, he decides that the bias is not so serious in the case of the arithmetic average with fixed weights, or aggregative index (which he designates as Formula 53) for, on page 219, he says: "There are two, [index numbers] 53 and 54, which, while free from bias, are not free from joint error." On page 361, he lists Formula 53 among "the eight most practical formulæ," and, on page 362, he advises that, between Census dates, the United States Bureau of Labor Statistics continue to use this formula in computing the index of wholesale prices.

On the other hand, he denounces the simple arithmetic average of relatives in no uncertain terms, for, on page 328,

he asserts that, "The simple arithmetic index number .... ranks among the 'worthless' formulæ in accuracy." On page 361, he continues: "The simple arithmetic (Formula 1) should not be used under any circumstances, being always biased and usually freakish as well."

We have seen that the geometric average is always smaller than the arithmetic average. Is it also true that, for some reason inherent in the formula, the trend of the weighted arithmetic average always differs from the trend of the unweighted arithmetic average? The facts are easily demonstrated by experimentation with hypothetical figures. Table 29 furnishes the necessary illustration.

TABLE 29

EFFECT OF A CHANGE IN WEIGHTS UPON THE TREND OF AN AVERAGE OF RELATIVES

TYPE OF WEIGHTING	WEIGHT	1925		1926		1927	
		Relative	Product	Relative	Product	Relative	Product
Equal weights (unweighted):							
Wheat	1	100	100	80	80	70	70
Corn	1	100	100	105	105	110	110
Oats	1	100	100	115	115	120	120
Total	3	.....	300	.....	300	.....	300
Average	.....	.....	100	.....	100	.....	100
Wheat weighted heavily; oats lightly:							
Wheat	4	100	400	80	320	70	280
Corn	2	100	200	105	210	110	220
Oats	1	100	100	115	115	120	120
Total	7	.....	700	.....	645	.....	620
Average	.....	.....	100	.....	92	.....	89
Wheat weighted lightly; oats heavily:							
Wheat	1	100	100	80	80	70	70
Corn	2	100	200	105	210	110	220
Oats	4	100	400	115	460	120	480
Total	7	.....	700	.....	750	.....	770
Average	.....	.....	100	.....	107	.....	110

Trend of weighted average of relatives dependent upon weights used.—The reader will observe that, in Table 29, the same relative prices are weighted in three different ways. When the prices of the three grains are given equal weights —that is when the average calculated is the simple or unweighted one—the trend is horizontal. However, when wheat, the falling commodity, is weighted heavily the trend of the averages declines. On the other hand, as might be expected, when oats, the rising commodity, receives the heavy weight, the trend of the averages is upward. Clearly, then, whether the trend of the weighted arithmetic averages or of the simple arithmetic averages is steeper depends wholly upon whether the rising or the falling prices are receiving the heavier weights.

We have seen, however, that Professor Irving Fisher's experiments with relatives of wholesale prices led him to conclude that such averages, when weighted by the values of the products in the base year, have a trend lying below the unweighted averages of the same items. If this relationship actually holds, it must be because of some tendency for the items which are weighted heavily in the base year to have declining tendencies while those that are weighted lightly in the base year have rising tendencies. One would hardly expect that, at all times, important articles are increasing in value less rapidly than unimportant articles, hence the question arises as to whether the phenomenon discovered by Professor Fisher is not merely a peculiarity arising from the characteristics of the particular year chosen as a base, the particular period covered, or the particular list of commodities which he selected for study. The only way to answer this question is to continue our experimentation, using, however, as raw material in the new test, prices of commodities at wholesale instead of the prices of shares of stock. The character of the data used in this last-mentioned experiment has been described in the immediately preceding chapter.

For the present experiment, six samples have been used. These six samples, 1A, 1B, 1C, 1D, 1E, and 1F, each consisting of 31 items, have been selected carefully on a proportional basis, from among the 336 cards bearing the respective records of price data for the various commodities.

The procedure used in making the selection was as follows:

I. The commodities were classified into the following major groups, the number of commodities in each being indicated by the numbers following the items.
   (a) Foodstuffs, 87.
   (b) Other goods sold to retailers to be resold for direct human consumption, 58.
   (c) Tools and materials used in building, 45.
   (d) Goods produced on farms or used primarily by farmers for indirect purposes, 42.
   (e) Goods for indirect use by the general public but not produced on farms, 104.

II. Each of these groups was subdivided into three classes based upon the importance assigned to the given article by the United States Bureau of Labor Statistics—an importance measured by the percentage which its value constituted of the total estimated value of all products in 1919, the base year. The three classes, respectively, include items rated individually at the following percentages of the total for all commodities:
   Class 1.—Under 0.10 per cent.
   Class 2.—0.10 and under 0.24 per cent.
   Class 3.—Over 0.24 per cent.

On this basis, the 336 commodities were divided into 15 classes containing from 8 to 46 items each. In choosing a sample of 31 items, the aim was to select from each class 1 item for each 11 items in the class, but always to include in the sample at least one item from each class. Each sample, then, when properly completed, contained 8 foodstuffs, 4 articles used in building, 4 farm products, 10 manufactured products, and 5 other commodities. Within the classes, the sampling was done at random.

Computation of indices of commodity prices.—Four separate averages of relatives were computed from the items con-

stituting each of the six samples selected. In calculating the weighted arithmetic average, the weights applied to the individual relatives were the percentages which the United States Bureau of Labor Statistics estimated the value of these particular commodities to constitute of the total value of all products in the United States in 1919. The following is an illustration of the appearance of a card containing all the data for a given commodity, when ready for combination with the data on other cards to form an index number.

Butter—Creamery: firsts, New York.

Percentage of value of all products represented by this commodity 0.37

Date	Price per Pound	Relative Price in 1919=100		Weight × Original Relative	Price Relative to Preceding Year
		Original	Log.		
Aug. 1915	$0.248	43.0	1.63347	15.91	
Aug. 1917	.405	70.2	1.84634	25.97	1.633
Av. 1919	.577	100.0	2.00000	37.00	1.424
May 1920	.591	102.4	2.01030	37.89	1.024
Jan. 1922	.344	59.6	1.77525	22.05	.582
Aug. 1925	.421	73.0	1.86332	27.01	1.224

The weighted arithmetic average of a given sample for a given date was computed by multiplying each price relative by its weight, summating the products on the 31 cards, and dividing the sum of the 31 products by the sum of the 31 weights. The other averages were computed by the processes previously described in connection with averages.

Behavior of indices derived by different formulæ.—Table 30 shows the index numbers derived by the application of four different formulæ to identical samples of commodity prices. One of the first things brought out by a study of this table is that these indices of commodity prices at whole-

sale show much less scatter than do the indices of prices of stocks, recorded in Table 26. The wide divergence in August, 1915, of the simple arithmetic average in Sample D is explained by the unusual behavior in that year of the price of phenol, an article for which the War caused a highly abnormal demand.

TABLE 30

A COMPARISON OF AVERAGES COMPUTED BY DIFFERENT AVERAGING PROCESSES FROM SIX PROPORTIONALLY SELECTED SAMPLES OF RELATIVES OF WHOLESALE PRICES

SAMPLES EACH CONTAINING 31 ITEMS	FORMULA APPLIED TO RELATIVES	INDICES OF WHOLESALE PRICES AT THE FOLLOWING DATES					
		Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
A	Simple arithmetic average	50.9	93.2	100.0	122.5	77.4	84.4
	Weighted arithmetic average	44.2	93.7	100.0	120.8	71.4	79.7
	Simple geometric average	48.9	86.1	100.0	118.8	72.8	82.5
	Median	46.0	78.7	100.0	117.8	75.8	82.8
B	Simple arithmetic average	49.0	85.0	100.0	127.9	75.9	83.5
	Weighted arithmetic average	50.5	86.6	100.0	150.1	74.9	82.4
	Simple geometric average	47.8	83.6	100.0	120.8	73.4	80.3
	Median	47.9	81.0	100.0	116.2	73.9	76.4
C	Simple arithmetic average	48.9	80.4	100.0	127.4	76.7	78.7
	Weighted arithmetic average	48.3	85.3	100.0	123.2	74.8	68.5
	Simple geometric average	46.3	78.0	100.0	125.0	74.3	74.8
	Median	47.6	73.8	100.0	130.3	75.9	73.5
D	Simple arithmetic average	92.0	106.4	100.0	123.6	76.8	87.5
	Weighted arithmetic average	66.9	108.2	100.0	130.9	79.7	87.0
	Simple geometric average	58.2	95.1	100.0	120.2	74.4	83.5
	Median	51.4	83.6	100.0	112.5	68.9	79.4
E	Simple arithmetic average	55.0	101.4	100.0	113.4	73.3	80.4
	Weighted arithmetic average	48.2	97.4	100.0	108.0	70.4	80.8
	Simple geometric average	49.8	91.9	100.0	111.1	68.5	77.7
	Median	49.1	73.1	100.0	109.3	70.9	77.3
F	Simple arithmetic average	47.4	94.5	100.0	117.8	67.5	76.1
	Weighted arithmetic average	51.5	96.9	100.0	119.5	73.6	75.2
	Simple geometric average	45.2	89.3	100.0	113.8	64.3	74.5
	Median	43.6	76.6	100.0	110.5	63.8	72.5

We know in advance that the simple geometric averages of relatives must, at every date except the base, be smaller

than the simple arithmetic averages of relatives, hence we are not surprised at the relative positions occupied by these two indices. We discover from the table that, in 22 cases out of 30, the medians of relatives are somewhat smaller than the geometric averages of the same. This is the reverse of the relationship found when these same formulæ were used to compute indices from actual share prices of stocks. Clearly, the nature of the data must be considered carefully in judging of the results to be expected from the use of any particular formula.

A comparison of weighted and unweighted arithmetic averages.—Because of Professor Fisher's findings concerning the relative positions of the simple and weighted arithmetic averages of relatives, especial interest attaches to the relationship of their trends as shown in Table 30. If we compare the extent of the rise in the two indices between August, 1915, and August, 1925, we find that, in three samples, the simple average showed the greater increase, while, in the other three samples, the weighted average rose more. If we add together the increases for all six samples we find that the total increase for the *weighted* average exceeded the total increase for the *simple* average by about 11 per cent. Clearly, these samples do not show any tendency for the simple average to climb above the weighted average.

Perhaps, however, we should expect to find a relationship similar to that of the geometric and arithmetic averages, and hence should look for a divergence from the base in both directions, with the simple arithmetic average always lying above the weighted. If we test the six samples recorded in Table 30 for such a relationship, we find, that the simple average was above the weighted one in the following number of instances:

Aug. 1915	Aug. 1917	May 1920	Jan. 1922	Aug. 1925
4	1	3	4	5

The simple average was, then, above the weighted average in 17 cases out of 30. These figures, on their face, indicate a slight tendency for the simple average to be the larger, but the excess is so small that it may well be due purely to chance. The evidence of Table 30 is so inconclusive that it has been felt worth while to extend the experiment by using 12 additional samples. To get this many samples, it was necessary to select six samples and make the computations; then to shuffle the cards in each class, and select six more samples; and finally, after re-shuffling to select six additional samples. The records for samples A, B, C, D, E, and F appear in Table 30. Those for samples k, l, m, n, o, p, q, r, s, t, u, and v are given in Table 31.

The increases between August, 1915, and August, 1925, recorded in Table 31 for the simple averages were, in 8 cases out of the 12, larger than those shown for the weighted averages. The findings here differ from those in Table 30 which showed a tie with three samples on each side.

When total points gained are compared, Table 31 shows an increase during the entire decade of 3,627 points for the simple averages as against 3,509 for the weighted averages, or a difference of 3 per cent. It will be remembered that the six samples in Table 30 showed a difference of 11 per cent in the opposite direction. With such diversity in findings it is clearly unsafe to generalize about differences in the trend of the weighted and unweighted averages of the wholesale prices in the United States during this particular decade.

A study of the 12 series of index numbers in Table 31 shows that, in the period before 1919, a period when prices rose enormously—the weighted averages were larger than the unweighted in 13 cases out of 24. Table 30 showed the weighted to be higher in 7 cases out of 12. On the face of the figures, in this period, the simple average tended to be

TABLE 31

A COMPARISON OF THE WEIGHTED AND UNWEIGHTED ARITHMETIC AVERAGES OF RELATIVE WHOLESALE PRICES DERIVED FROM TWELVE PROPORTIONALLY SELECTED SAMPLES

SAMPLES EACH CONTAINING 31 ITEMS	ARITHMETIC AVERAGE APPLIED TO RELATIVES	INDICES OF WHOLESALE PRICES AT THE FOLLOWING DATES					
		Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
k	Simple	54.5	87.4	100.0	119.7	76.9	78.3
	Weighted	52.0	82.5	100.0	116.5	76.5	75.6
l	Simple	48.3	80.7	100.0	124.6	73.0	82.6
	Weighted	45.0	85.1	100.0	127.5	74.9	82.0
m	Simple	48.1	85.5	100.0	129.3	75.8	82.8
	Weighted	51.4	84.8	100.0	119.1	88.6	85.1
n	Simple	55.0	90.9	100.0	120.1	75.8	80.1
	Weighted	53.4	94.2	100.0	124.3	73.8	81.7
o	Simple	48.4	83.5	100.0	121.2	77.6	83.8
	Weighted	47.4	86.3	100.0	101.2	64.1	78.3
p	Simple	53.1	93.6	100.0	125.7	71.0	78.7
	Weighted	49.4	91.8	100.0	147.3	68.6	72.1
q	Simple	53.8	95.7	100.0	119.6	74.4	81.8
	Weighted	44.3	94.3	100.0	117.8	68.6	74.6
r	Simple	50.0	88.4	100.0	123.9	74.1	80.7
	Weighted	52.7	85.5	100.0	144.7	70.2	79.6
s	Simple	57.2	91.9	100.0	117.5	72.1	79.9
	Weighted	58.4	97.2	100.0	114.9	65.6	77.3
t	Simple	44.6	88.0	100.0	124.5	69.9	75.2
	Weighted	45.0	105.2	100.0	117.8	64.5	74.8
u	Simple	50.1	86.0	100.0	119.5	79.4	82.0
	Weighted	48.3	88.0	100.0	116.5	79.0	81.9
v	Simple	47.9	86.4	100.0	140.6	76.7	87.8
	Weighted	52.7	83.8	100.0	161.7	72.8	87.9

lower than the weighted average—but the difference is so slight that it might well be due purely to chance.

The indices in Table 31, for dates after 1919, however, show the weighted average to be the larger in only 10 cases out of 36. In view of the fact that only 6 out of the 18 of the weighted indices in Table 30 were as large as the simple averages derived from the same samples, it appears that, between 1919 and 1925, there was an unmistakable tendency for the simple averages to decline less than the weighted, indicating that, in the price debâcle following 1919, the important commodities suffered most severely; though they had risen but little if any more than the unimportant commodities in the inflation period, 1915–1919. Clearly, it is unsafe to generalize concerning index numbers of wholesale prices by studying one period only, for conclusions drawn from one period may have no validity even in the years immediately following.

Simple and weighted arithmetic averages of relatives of various classes of commodities.—To understand better the reasons for the comparative behavior of the simple and weighted arithmetic averages of relatives of wholesale prices presented in Tables 30 and 31, two more Tables, namely, 32 and 33, have been constructed. These show, respectively, for each class of commodities, included in the entire list of 336 articles, the simple and weighted arithmetic averages of relatives.

A study of the data in these tables reveals the fact that, when all commodities are combined, the weighted averages rose somewhat more steeply than the simple averages between 1915 and 1919, and they underwent a slightly larger decline between 1919 and 1925. This confirms the evidence given by the index numbers derived from the 18 samples.

When we examine the different groups of commodities, we find that, during the rapid rise in prices from 1915 to 1919, the score for gain in altitude was 2 in favor of

TABLE 32

THE AVERAGE VALUES BY CLASSES OF COMMODITIES OF THOSE QUANTITIES OF THE RESPECTIVE ARTICLES SELLING ON THE AVERAGE FOR $1.00 IN 1919

PERCENTAGE OF TOTAL OUTPUT REPRESENTED BY INDIVIDUAL COMMODITIES	NUMBER OF ITEMS IN CLASS	UNWEIGHTED AVERAGE VALUE FOR CLASS AT THE FOLLOWING DATES					
		Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
All classes of goods:							
Total	336	54.62	92.51	100.00	121.94	74.28	81.76
Under .10	126	61.71	96.11	100.00	119.67	75.44	83.06
.10 to .24	93	51.45	87.47	100.00	122.66	72.78	81.35
.24 and over	117	49.51	92.63	100.00	123.81	74.23	80.68
Foodstuffs:							
Total	87	49.15	82.45	100.00	104.07	66.87	79.25
Under .10	31	50.49	81.42	100.00	98.76	63.78	74.18
.10 to .24	25	47.88	84.56	100.00	103.35	66.78	79.73
.24 and over	31	48.83	81.79	100.00	109.95	70.05	83.92
Other goods sold to retailers to be resold for direct human consumption:							
Total	58	50.78	75.61	100.00	129.47	93.06	92.34
Under .10	14	47.19	78.69	100.00	123.81	97.26	98.04
.10 to .24	23	48.59	72.71	100.00	135.64	86.50	85.97
.24 and over	21	55.58	76.72	100.00	126.48	97.45	95.53
Tools and materials used in building:							
Total	45	47.60	75.35	100.00	139.80	91.77	87.73
Under .10	27	48.31	78.04	100.00	130.83	91.37	89.27
.10 to .24	10	44.39	70.89	100.00	141.60	87.75	85.98
.24 and over	8	49.20	71.85	100.00	167.81	98.16	84.74
Goods produced on farms or used primarily by farmers for indirect purposes:							
Total	42	47.10	86.47	100.00	113.86	53.93	63.12
Under .10	8	47.64	86.19	100.00	127.34	61.72	62.44
.10 to .24	12	49.07	85.94	100.00	115.42	57.47	64.59
.24 and over	22	45.83	86.87	100.00	108.11	49.17	62.56
Goods for indirect use by the general public but not produced on farms:							
Total	104	67.42	120.20	100.00	128.23	70.66	82.89
Under .10	46	84.01	123.63	100.00	124.60	69.70	84.42
.10 to .24	23	62.50	113.40	100.00	126.23	67.04	85.23
.24 and over	35	48.85	120.15	100.00	134.28	74.29	79.35

TABLE 33

WEIGHTED RELATIVE PRICES OF DIFFERENT TYPES OF GOODS THE WEIGHTS BEING THE ESTIMATED VALUES OF THE TOTAL OUTPUTS OF THE RESPECTIVE ARTICLES IN 1919

PERCENTAGE OF TOTAL OUTPUT REPRESENTED BY INDIVIDUAL COMMODITIES	NUMBER OF COMMODITIES IN CLASS	PERCENTAGE OF TOTAL OUTPUT REPRESENTED BY CLASS	WEIGHTED AVERAGE RELATIVE PRICES AT DATES GIVEN					
			Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
All classes of goods:								
Total	336	98.92*	50.75	91.86	100.00	123.96	73.11	80.11
Under .10	126	5.63	65.80	98.42	100.00	121.52	75.09	83.49
.10 to .24	93	15.30	52.01	86.98	100.00	122.37	72.56	81.95
.24 and over	117	77.99	49.41	92.34	100.00	124.45	73.08	79.51
Foodstuffs:								
Total	87	22.86	50.58	83.80	100.00	118.15	69.12	81.27
Under .10	31	1.45	51.41	84.32	100.00	103.97	65.03	75.45
.10 to .24	25	4.33	48.53	83.80	100.00	102.90	67.16	81.12
.24 and over	31	17.08	51.03	83.75	100.00	123.22	69.97	81.80
Other goods sold to retailers to be resold for direct human consumption:								
Total	58	18.73	55.71	79.73	100.00	125.83	99.70	94.73
Under .10	14	0.47	50.72	82.53	100.00	117.66	101.72	100.21
.10 to .24	23	3.36	48.30	72.89	100.00	135.70	86.77	87.30
.24 and over	21	14.90	57.54	81.19	100.00	123.87	102.55	96.23
Tools and materials used in building:								
Total	45	6.20	47.42	72.66	100.00	153.71	94.00	86.07
Under .10	27	1.33	48.15	76.70	100.00	135.84	90.62	88.26
.10 to .24	10	1.65	43.76	69.56	100.00	144.45	87.91	85.39
.24 and over	8	3.22	49.00	72.57	100.00	165.84	98.52	85.51
Goods produced on farms or used primarily by farmers for indirect purposes:								
Total	42	25.12	43.15	85.50	100.00	105.40	51.29	66.30
Under .10	8	0.39	57.10	107.95	100.00	110.97	67.00	71.08
.10 to .24	12	2.14	49.24	83.79	100.00	115.90	57.41	64.31
.24 and over	22	22.59	42.33	85.28	100.00	104.31	50.44	66.41
Goods for indirect use by the general public but not produced on farms:								
Total	104	26.01	55.44	118.39	100.00	138.54	73.58	80.50
Under .10	46	1.99	93.34	125.10	100.00	127.70	67.35	84.64
.10 to .24	23	3.82	64.34	112.28	100.00	126.80	68.05	86.59
.24 and over	35	20.20	50.03	118.88	100.00	141.83	75.24	78.94

* Not all of the commodities entering into the United States Bureau of Labor Statistics price index are included here; hence the total is somewhat less than 100 per cent.

weighted, 2 in favor of unweighted, and a tie in the fifth case. During the price decline after 1919, the unweighted averages fell more in 3 cases and the weighted fell more in 2 cases. We find further that, between 1915 and 1919, the unimportant commodities rose in price less than the important commodities. On the other hand, the unimportant commodities declined slightly less between 1919 and 1925 than did the important commodities.

Differences in trends of weighted and unweighted arithmetic averages of relatives due to chance variations.—All in all, the evidence seems to make it perfectly clear that it is wholly unsafe to conclude that there is any relationship existing continuously over any considerable period of time between the importance of a commodity and either the direction or the extent of its price movement. If this is true, it follows that there is no ground for assuming that the simple arithmetic average of relatives tends, in general, to be larger or to have a steeper upward trend than does the weighted arithmetic average of relatives. Since there is apparently insufficient evidence to show any such divergence in trends (or so-called bias), manifestly all conclusions built upon this hypothesis likewise fall to the ground.

Summary

1. The behavior of an index number is easier to interpret if one is familiar with the characteristics of the type or average used in the computation of the index number.

2. When numbers of different sizes are averaged, the arithmetic average is always larger and the harmonic average always smaller than the geometric average. This fact explains why averages of relatives of these three types, when plotted, diverge from the base with the arithmetic average at the top and the harmonic average at the bottom. They

are brought together at the base by the fact that, at that point, all the relatives are of the same size.

3. When original data (instead of relatives) are averaged, there is no simple rule which determines in advance whether the application of the arithmetic, the geometric, or the harmonic average will give the series of index numbers having the highest trend.

4. A peculiarity of a series of geometric averages of relatives is that their relationships to each other are not changed by shifting the base used for the individual relatives.

The importance of this characteristic has been greatly over-emphasized; but it may make the geometric average a convenient one to use in certain cases.

5. Reducing numbers to relatives on a given base is equivalent to assigning to each series a specific set of weights.

6. A study of three different index numbers of stock prices, computed from samples each containing 31 items and each using prices of November, 1916, as a base, shows the steepness of the trends of relatives between November, 1916, and March, 1925, to have ranked in the following sequence, the steepest being placed first:

(a) Arithmetic means of relatives.
(b) Relatives of aggregates of actual prices per share.
(c) Medians of actual prices per share.
(d) Medians of relatives.
(e) Geometric averages of actual prices per share.

The dispersion between these index numbers is quite large. The difference in the trends is due in the main, to the fact that, during this particular period, the high-priced stocks rose much more slowly than did the stocks which were low at the beginning of the period.

7. Medians and geometric averages are less affected than are arithmetic averages by extreme changes in individual items entering into the average.

8. The greater the dispersion in the data averaged, the greater the divergence between the arithmetic and the geometric averages.

9. A study of records of the prices of goods at wholesale between August, 1915, and August, 1925, shows

(a) That changing the formula used in the computation of the index numbers gives rise to much less dispersion among indices of wholesale prices than it does among indices of stock prices.

(b) That, contrary to the contention of Professor Fisher, there is no general tendency for the simple arithmetic average of relatives to show a trend differing from that of the weighted arithmetic average. Such differences as appear, arise solely from peculiarities of the data averaged. Whenever important commodities rise faster than unimportant commodities, the trend of the weighted arithmetic averages of relatives will be above that of the simple arithmetic averages of the same—and *vice versa.* It appears, then, that there is no foundation for the belief that the unweighted arithmetic average of relatives has more of a "bias" than the weighted arithmetic average of relatives.

QUESTIONS AND PROBLEMS

1. Which of the following types and averages are not applicable to a group containing negative items:

(a) Arithmetic?
(b) Harmonic?
(c) Geometric?
(d) Median?
(e) Mode?

2. Is an index series composed of medians of relatives better adapted to periods in which extremely erratic price movements occur or to a period in which great irregularities are practically absent?

3. In preparing an index series, why is it frequently considered preferable to use the average of the three or five central items in an array instead of using the median item only?

4. Which average of relatives is not affected by changing the base of the relatives? Does this peculiarity give to it any especial fitness for use in the computation of index numbers?

5. In the case of what pairs of the averages listed in Question 1 does the difference vary directly with the dispersion of the numbers averaged?

6. Does the fact that the arithmetic average is larger than the geometric average justify one in ascribing an "upward bias" to index numbers obtained by taking the arithmetic average of relatives?

7. In November, 1916, was it more profitable to invest in high-priced or in low-priced shares of stock? Do you think that it always pays to buy low-priced shares? Is your answer borne out by a careful analysis of the figures in Table 28?

8. If, in a given period, the prices of important commodities rose relatively faster than the prices of unimportant commodities, would a simple or a weighted arithmetic average of relatives show the sharper upward trend?

9. Is it true that the use of the simple arithmetic average of relatives in the computation of an index series tends to give to that series a sharper upward trend than that of a series derived by using a weighted arithmetic average of relatives with weights kept constant?

BIBLIOGRAPHICAL NOTE

Most text books on statistics treat of the different averages and their characteristics. Their formulæ are also compared by Irving Fisher in *The Making of Index Numbers,* 1922, pp. 27–40.

The supposed merits of the different averages in the computation of index numbers are discussed by most writers on index numbers, but the material along this line is voluminous rather than enlightening. The article by Frederick R. Macaulay in *The Journal of the American Statistical Association,* March, 1926, pp. 27–39, is, however, worth reading.

Concrete illustrations of the varying results obtained by applying different averages to the same data are presented by the following writers:

G. H. Knibbs, Australian Commonwealth Bureau of Labor and Statistics, *Report No. 1,* December, 1912, Appendix VIII, pp. xxiii–xxxviii.

Wesley C. Mitchell, *Journal of Political Economy,* July, 1916, pp. 626–651, and United States Bureau of Labor Statistics, *Bulletin 284* (1921), pp. 11–23, 33–59, and 94–105.

Irving Fisher, *The Making of Index Numbers* (1922), entire book.

Norman Crump, *Journal of the Royal Statistical Society,* March, 1924, p. 195.

L. H. Bean and O. C. Stine, *Journal of the American Statistical Association,* March, 1924, pp. 30–35.

The effect upon index numbers of giving different weights to the different items included has been touched upon by many writers. The following are a few of the numerous worth while discussions:

F. Y. Edgeworth, *Journal of the Royal Statistical Society,* December, 1883, pp. 714–718.

Correa M. Walsh, *The Measurement of General Exchange Value* (1901), pp. 80–97.

H. Fountain, *Memorandum on the Construction of Index Numbers of Prices, House of Commons Paper, No. 321* (1903), pp. 429–449.

A. L. Bowley, *Elements of Statistics* (1907), pp. 111–118 and *Journal of the Royal Statistical Society,* March, 1926, pp. 300–319.

G. H. Knibbs, Australian Commonwealth Bureau of Labor and Statistics, *Report No. 1,* December, 1912. Appendix VIII, pp. xxiii–xxxviii.

A. C. Pigou, *The Economics of Welfare* (1920), pp. 69–90.

Wesley C. Mitchell, *Journal of Political Economy,* July, 1916, pp. 684–691, and United States Bureau of Labor Statistics, *Bulletin 284* (1921), pp. 59–68.

Irving Fisher, *The Making of Index Numbers* (1922), pp. 439–450.

Forrest B. Ashby, *Journal of the American Statistical Association,* March, 1928, pp. 49–54.

Woodlief Thomas, *Journal of the American Statistical Association,* March, 1928, *Papers and Proceedings,* pp. 95–97.

CHAPTER VII

PERCENTAGES OF ERROR FOUND IN CERTAIN PRICE INDICES

Experimentation necessarily limited in scope.—The preceding discussion, if it has succeeded in its purpose, has cleared the ground for what is presumably the most important part of the experimental section of this study—namely, the comparison, by means of actual measurements, of the errors found in a few of the most common types of index numbers when they are used to answer three questions—one concerning stock prices and two concerning commodity prices.

The reader will, of course, inquire why experimental work has been confined to a study of three questions and such a limited number of types of index numbers when there are an infinite number of questions to be answered and when so many highly recommended formulæ might be tested. The obvious answer is that, even in the case of a single question, the cost of testing all the possible formulæ and publishing the results would be prohibitive. All that is attempted here is to indicate the method of attack and to present the results for a few formulæ. However, these formulæ are the ones actually used in the computation of the large majority of index numbers published at present; hence, the usefulness of the findings is not so limited as one might suppose. The mistake must not be made, however, of assuming that the percentages of error found in the index numbers here are representative of errors in the same index numbers when

applied to other classes of data or even when applied to the same type of data under conditions giving rise to price relationships radically different from those prevailing during the period studied.

Established facts concerning index numbers.—At the risk of boring the reader, we will here reiterate certain of the fundamental principles previously discussed. They are:

1. There is one formula, and one formula only, which, if applied to complete data, will answer any specific question.
2. Index numbers are devices computed from sample data which are intended to approximate as closely as possible the facts which would be revealed by a study of complete data.
3. That index series is most meritorious which deviates least from the course of the relatives representing variations in the totality of items.
4. The index number which comes nearest to answering a specific question may come very far from answering a slightly different question. Obviously, then, there can be no such thing as an "ideal" index number.

Formula applicable to totality commonly assumed to be best for use with sample data.—Even though there is no "ideal" index number for all purposes, it may, nevertheless, still be true that there is a specific formula for index number construction which gives best results in answering a certain type of question pertaining to a certain type of data. With complete data, the exact formula necessary to answer the question is determined by the nature of the question. The most natural assumption—and the one which seems to have been tacitly accepted by the great majority of authorities on index numbers—is that this same formula, when applied to sample data, will produce the index series showing the least deviation from the course of the relatives portraying the fluctuation in the totality. Since we have already discovered that several of the other commonly accepted premises concerning index numbers are wholly erroneous, it may be worth while to examine this one carefully before accepting it as an established fact.

STOCK-PRICES PER SHARE

Nature of stock-price data used in experiments.—The records of stock prices described in earlier chapters furnish an admirable approximation to a specific totality of items for use in experimentation, for the 302 issues, aggregates for which are included in Table 27, cover practically all the important stocks listed on the leading stock exchanges in the United States, which were quoted at each of the different dates covered. Since we know exactly what variations occurred in the prices of each of these stocks, we are in a position to give exact answers to certain questions concerning the relative movements of the prices of the 302 stocks taken together. We can also experiment freely with various averages of the sample stocks selected from the list to see how the variations in these averages compare with the variations in the averages of the totality.

Average of totality not identical with average of universe.—Attention should, however, be called to the fact that the errors, as measured from the *totality* shown in Table 27, will not be exactly the same as would errors measured from relatives representing an ideal universe including at least the prices of all stocks existing in the United States and perhaps those of imaginary stocks that might be thought of as existing under the economic conditions prevailing. As to just what the differences in percentages of error would be if we had such an ideal universe as a point of departure, it is manifestly impossible to say. Presumably, in the case of small samples, the discrepancies would, in most cases, be unimportant. However, whenever the number of items in the sample is so increased that the size of the sample begins to approach the totality of 302 stocks, and whenever the same formula is applied both to the sample and the totality, it is evident that the computed error will be smaller than the true error—namely, than the difference between

the average of the sample and the average of an ideal universe including every stock sold. Nevertheless, a comparison of the index numbers derived from the various samples with the relatives representing the variations in the totality will constitute the best tests yet made of the usefulness of the various kinds of index numbers for answering the particular question under consideration.

Clearly, in making such tests, we must deal with one question at a time, and each question must be precisely stated. Our first question deals with the deviations of the trend of the totals of the sample from the trend of the totals of the prices per share of all 302 stocks.

DEVIATIONS FROM THE AGGREGATE OBTAINED BY ADDING THE PRICES PER SHARE OF 302 STOCKS

QUESTION 1

What would have been the relative total market values, at the seven dates specified, of the stocks possessed by a man who invested in November, 1916, in one share of each of the 302 important industrial stocks (data for which are presented in Table 27) had he retained all stocks received as dividends thereon, and sold none of his holdings before April, 1925?

With complete data questions can be answered precisely. —Aside from possible errors arising either from the necessity, in a few instances, of making estimates of prices, or from slips in computation, the question is answered absolutely by the figures appearing in the first line of Table 27. The value of the holdings at each date, relative to the November, 1916, high, is shown in the first line of Table 28. The relative comparison answering the question has been obtained from complete data and the mode of obtaining the figures is not open to controversy. There exists but one correct method—the one used.

Reason for using samples instead of complete data.—To collect the information necessary to answer this question for only seven dates has required, however, a very considerable amount of labor. To extend the study to all the months for even one decade would involve a large amount of effort. Is it possible to approximate the same results by using a small but carefully selected sample to represent the whole? If this device can be successfully utilized, a large proportion of the work involved in the making of comparisons may be eliminated. We have here a typical example illustrating the problem of index numbers.

Testing index numbers for accuracy.—It is obvious, as has been previously stated, that, if the *same* formula is applied to a sample that is used with the totality which it is intended to represent, the difference between the average of the sample and the average of the totality will tend to grow less, as the size of the sample is increased, until, when the size of the sample equals the totality, the difference will be reduced to zero. If, however, the mathematical process used with the sample is *different* from that applied to the totality, the probabilities are that a discrepancy between the two results will persist—no matter what percentage of the items is included in the sample. With some pairs of formulæ, differing greatly in their make-up, this ultimate difference is likely to be very large. The fact just stated necessarily inclines us to expect that the formula which answers the question with complete data will, when applied to a small sample, give more accurate index numbers than can be obtained by any other method. As before stated, however, this hypothesis stands in need of proof. To test its validity, 7 different kinds of averages have been computed for each of the samples—1a, 1b, and 1c—samples, the selection of which was described in some detail in an earlier chapter.

An index series has been derived from each of the three

TABLE 34

THE DEVIATIONS OF DIFFERENT INDEX NUMBERS OF STOCK PRICES FROM THE ACTUAL RELATIVES ANSWERING QUESTION 1 (PROPORTIONAL SAMPLES)

Relatives Computed from	Sample: Number	Sample: Number of Items	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High	Average of Deviations for 6 Dates
Aggregates:										
Index	1a	30	100.00	72.93	104.92	61.50	92.96	74.11	105.70	
	1b	30	100.00	73.22	115.49	69.97	113.68	97.86	135.89	
	1c	30	100.00	68.05	110.40	64.36	99.69	85.83	95.49	
Deviation	1a	.....	.....	1.32	2.93	4.50	11.24	10.15	5.45	5.93
	1b	.....	.....	1.61	7.64	3.97	9.48	13.60	24.74	10.17
	1c	.....	.....	3.56	2.55	1.64	4.51	1.57	15.66	4.92
	Av. Dev.	.....	.....	2.16	4.37	3.37	8.41	8.44	15.28	7.01
Sum of 2 median items:										
Index	1a	30	100.00	63.73	102.91	36.75	72.70	49.60	77.39	
	1b	30	100.00	74.42	109.66	63.13	99.80	87.62	101.97	
	1c	30	100.00	66.90	115.24	56.04	99.74	78.02	119.21	
Deviation	1a	.....	.....	7.88	4.94	29.25	31.50	34.66	33.76	23.67
	1b	.....	.....	2.81	1.81	2.87	4.40	3.36	9.18	4.07
	1c	.....	.....	4.71	7.39	9.96	4.46	6.24	8.06	6.80
	Av. Dev.	.....	.....	5.13	4.71	14.02	13.45	14.75	17.00	11.51
Sum of 6 median items:										
Index	1a	30	100.00	64.84	101.27	43.94	77.56	52.28	83.88	
	1b	30	100.00	71.45	109.90	67.11	99.27	88.58	102.40	
	1c	30	100.00	68.50	108.67	57.44	95.00	83.47	113.85	
Deviation	1a	.....	.....	6.77	6.58	22.06	26.64	31.98	27.27	10.22
	1b	.....	.....	.16	2.95	1.11	4.93	4.32	8.75	3.55
	1c	.....	.....	3.11	0.82	8.56	9.20	0.79	2.70	4.20
	Av. Dev.	.....	.....	3.35	3.15	10.58	13.59	12.36	12.91	9.32

Sum of 2 median relatives:										
Index	1a	30	100.00	64.60	93.80	60.65	87.50	74.65	84.80	
	1b	30	100.00	72.55	98.85	71.60	87.35	81.85	90.90	
	1c	30	100.00	66.90	91.50	66.80	89.90	77.85	95.80	
Deviation	1a	.....	.....	7.01	14.05	5.35	16.70	9.61	26.35	13.18
	1b	.....	.....	.94	9.00	5.60	16.85	2.41	20.25	9.17
	1c	.....	.....	4.71	16.35	.80	14.30	6.41	15.35	9.65
	Av. Dev.	.....	.....	4.22	13.13	3.92	15.95	6.14	20.65	10.67
Mean of relatives:										
Index	1a	30	100.00	66.70	119.46	63.19	101.13	77.52	128.99	
	1b	30	100.00	84.34	174.63	100.87	201.84	188.06	355.95	
	1c	30	100.00	65.51	113.08	63.43	98.15	78.93	119.10	
Deviation	1a	.....	.....	4.91	11.61	2.81	3.07	6.74	17.84	7.83
	1b	.....	.....	12.73	66.78	34.87	97.64	103.80	244.80	93.44
	1c	.....	.....	6.10	5.23	2.57	6.05	5.33	7.95	5.54
	Av. Dev.	.....	.....	7.91	27.87	13.42	35.59	38.62	90.20	35.60
Aggregate of all items except 4 extremes:										
Index	1a	30	100.00	71.61	103.86	63.77	85.29	66.23	95.67	
	1b	30	100.00	75.16	121.10	72.72	108.98	93.34	109.70	
	1c	30	100.00	70.52	102.76	63.80	93.20	77.90	68.09	
Deviation	1a	.....	.....	0.00	3.99	2.23	18.91	18.03	15.48	9.77
	1b	.....	.....	3.55	13.25	6.72	4.78	9.08	1.45	6.47
	1c	.....	.....	1.09	5.09	2.20	11.00	6.36	43.15	11.48
	Av. Dev.	.....	.....	1.55	7.44	3.72	11.56	11.16	20.03	9.24
Geometric average of original items:										
Index	1a	30	100.00	61.36	97.26	50.42	78.79	50.84	76.46	
	1b	30	100.00	69.36	123.79	64.72	102.59	63.67	90.05	
	1c	30	100.00	59.92	96.36	49.37	75.12	49.41	68.88	
Deviation	1a	.....	.....	10.25	10.59	15.58	25.41	33.42	34.69	21.66
	1b	.....	.....	2.25	15.94	1.28	1.61	20.59	21.10	10.46
	1c	.....	.....	11.69	11.49	16.63	29.08	34.85	42.27	24.34
	Av. Dev.	.....	.....	8.06	12.67	11.16	18.70	29.62	32.69	18.82
Aggregate of all companies:										
Relatives	.....	302	100.00	71.61	107.85	66.00	104.20	84.26	111.15	

samples, using as a base in every case the "high" of November, 1916. For each of the other six dates covered in the study, the deviations of the respective index numbers from the relative of the totality (also computed on the base November, 1916) have been computed. These deviations, and the averages thereof, are recorded in Table 34.

The relative merits of the indices tested.—From the evidence thus far secured, it appears that the expectation that the best results would be secured by employing the formula used with the totality has been fulfilled, for the first set of relatives derived from the aggregates of the samples (the formula used with the totality), shows a distinctly smaller average error than is obtained with any of the other averages or types. For answering this question, the mean of the relatives is the worst of the seven averages tested, and the geometric average of the original items shows the next largest error. All of the other errors are distinctly larger than those obtained by using relatives of aggregates. However, the "Sum of the 6 Median Items," and the "Aggregate of All Items Except the 4 Extremes," give moderately satisfactory results; but, even so, more accurate index numbers are secured by including than by excluding the extreme items in the sample.

Increase of error with departure from base.—An interesting phenomenon is revealed by Table 34—a phenomenon which, so far as the writer is aware, has not previously been discovered—namely, that the percentage of error tends to increase rapidly as the time-distance from the base is increased. This increase is presumably due to the fact that the prices of certain stocks have marked upward or downward trends, these trends depending upon changing market conditions or the skill of the management. A tendency, even in a single stock, to move sharply either in an upward or in a downward direction, is sufficient to make the average of the sample veer in the same direction and to cause the error

to increase with the passage of time. The fact, that, in the case of stock prices, there is clearly evident a tendency for the error to increase with the time-distance from the base, does not justify the belief that the same tendency is also inherent in other types of price data. The facts concerning other fields can only be determined experimentally.[1]

Relation between the size of a sample and its percentage of error.—Manifestly, if one applies to the sample the same formula used with the total of 302 stocks, the average of the sample will tend, as the size of the sample is increased, to approach that of the totality. While this approach will, of course, be irregular, yet the general tendency would be expected to approximate a smooth curve representing some mathematical equation. An effort has been made to determine empirically the general nature of this curve.

In the preceding chapter, a description was given of the method of selecting the three samples of 30 stocks each labelled 1a, 1b, and 1c. This same method was later extended to four other samples of the same size. These additional samples have been designated 1d, 1e, 1f, and 1g, respectively. To lessen the amount of labor involved in computing averages for samples of large size, the totals of these seven small samples have been combined in a variety of ways, to give totals for larger samples. By combinations of this sort, 31 samples have been secured, there being 5 samples in each of 6 groups, and 1 in the last group. The samples in the first group contain 30 stocks each. In the succeeding groups they include 60, 90, 120, 150, 180, and 210 stocks, respectively.

[1] A related phenomenon in the field of commodity prices at wholesale is discussed by Dr. Frederick C. Mills on pages 256–286 of the report of The National Bureau of Economic Research entitled *The Behavior of Prices*. Since, however, his indices pertain to commodity prices and are computed from price relatives, their relationship to the facts revealed in this study of absolute prices of stocks is not very close.

The average of each sample for each date has been computed; the averages thus obtained have been reduced to index numbers on the base November, 1916; the deviation of the index for each date from the relative of the totality at the same date has been ascertained; these average deviations have been expressed as percentages of the corresponding relative representing the totality; and the percentages for the five samples have been averaged for each group except the last. This rather complicated process is illustrated in Tables 35, 36, and 37.

It will be noted that, in Table 37, both the arithmetic average of the percentages of error and the median of the percentages of error have been computed for each size of sample at each date. An inspection of the results indicates that the arithmetic averages tend to diminish in size with the increase in the size of the sample in a manner more regular than that characterizing the medians. All further analyses of the data are, therefore, based upon the arithmetic averages rather than upon the medians.

The figures in Table 37 reveal what one would naturally anticipate—namely, that, as more stocks are taken into the sample, the influence of any one stock is lessened, and errors in opposite directions begin to counterbalance each other. We see also that, as the size of the sample is increased, the percentage of error in the index number shows a less marked tendency toward growth in the later periods. This must necessarily be true, since, when all the items are included in the sample, the error must entirely disappear, no matter what the nature of the data considered.

Cases are not lacking, however, in which the influence of the time-distance from the base has been so important as to overshadow the effect of varying the size of the sample. For example, the figures in Table 37 for December, 1917, show that a sample of 30 items is large enough to assure, in most cases, an error of less than 5 per cent. For August,

TABLE 35

INDEX NUMBERS DERIVED BY THE METHOD OF AGGREGATES FROM DIFFERENT-SIZED SAMPLES OF PRICES OF STOCKS OF CORPORATIONS

SAMPLE		ORIGINAL INDICES						
Composition	Number of Items	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High
1a	30	100.00	72.93	104.92	61.50	92.96	74.11	105.70
1b	30	100.00	73.22	115.49	69.97	113.68	97.86	135.89
1c	30	100.00	68.05	110.40	64.36	99.69	85.83	95.49
1d	30	100.00	73.06	109.65	63.08	93.85	81.80	102.01
1e	30	100.00	74.19	108.50	56.37	89.27	62.47	90.15
1a+1b	60	100.00	73.09	110.65	66.09	104.19	86.98	122.06
1b+1c	60	100.00	70.70	113.00	67.23	106.84	91.98	116.15
1c+1d	60	100.00	70.63	110.02	63.70	96.69	83.76	98.84
1d+1e	60	100.00	73.71	108.99	59.21	91.21	70.66	95.17
1e+1a	60	100.00	73.71	107.14	58.32	90.67	66.90	96.06
1a+1b+1c	90	100.00	71.37	110.57	65.50	102.65	86.59	112.99
1d+1e+1a	90	100.00	73.51	107.93	59.81	91.67	71.56	97.92
1b+1c+1d	90	100.00	71.50	111.86	65.81	102.41	88.51	111.32
1e+1a+1b	90	100.00	73.56	109.73	61.94	97.82	76.51	108.43
1c+1d+1e	90	100.00	72.09	109.39	60.68	93.63	75.00	95.26
1a+1b+1c+1d	120	100.00	71.82	110.32	64.86	100.32	85.32	110.08
1b+1c+1d+1e	120	100.00	72.35	110.80	62.82	98.25	80.26	104.62
1c+1d+1e+1a	120	100.00	72.26	108.49	60.85	93.50	74.82	97.37
1d+1e+1a+1b	120	100.00	73.44	109.71	62.21	96.87	77.77	106.89
1e+1a+1b+1c	120	100.00	72.30	109.89	62.49	98.24	78.64	105.47
1a+1b+1c+1d+1e	150	100.00	72.45	109.83	62.60	97.38	79.24	104.77
1b+1c+1d+1e+1f	150	100.00	71.96	109.16	63.05	98.32	80.72	104.68
1c+1d+1e+1f+1g	150	100.00	72.50	107.94	65.17	100.54	81.44	102.35
1d+1e+1f+1g+1a	150	100.00	73.33	107.04	64.75	99.54	79.57	104.03
1e+1f+1g+1a+1b	150	100.00	73.36	108.10	66.00	103.13	82.45	110.16
1f+1g+1a+1b+1c+1d	180	100.00	72.30	108.62	67.43	104.18	87.36	111.02
1g+1a+1b+1c+1d+1e	180	100.00	73.02	109.60	65.64	101.93	82.82	107.55
1a+1b+1c+1d+1e+1f	180	100.00	72.08	108.58	62.84	97.59	79.83	104.81
1b+1c+1d+1e+1f+1g	180	100.00	72.61	109.07	65.90	102.52	83.90	107.38
1c+1d+1e+1f+1g+1a	180	100.00	72.55	107.53	64.70	99.56	80.49	102.77
1a+1b+1c+1d+1e+1f+1g	210	100.00	72.64	108.61	65.40	101.44	82.79	107.20
All stocks	302	100.00	71.61	107.85	66.00	104.20	84.26	111.15

TABLE 36

ERRORS OF STOCK-PRICE INDEX NUMBERS DERIVED FROM SAMPLES OF DIFFERENT SIZES BY METHOD OF AGGREGATES

SAMPLE		ACTUAL DEVIATIONS FROM RELATIVES REPRESENTING THE TOTALITY						
Composition	Number of Items	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High
1a.....	30	0	1.32	2.93	4.50	11.24	10.15	5.45
1b.....	30	0	1.61	7.64	3.97	9.48	13.60	24.74
1c.....	30	0	3.56	2.55	1.64	4.51	1.57	15.66
1d.....	30	0	1.45	1.80	2.92	10.35	2.46	9.14
1e.....	30	0	2.58	0.65	9.63	14.93	21.79	21.00
1a+1b.....	60	0	1.48	2.80	0.09	0.01	2.72	10.91
1b+1c.....	60	0	0.91	5.15	1.23	2.64	7.72	5.00
1c+1d.....	60	0	0.98	2.17	2.30	7.51	0.50	12.31
1d+1e.....	60	0	2.10	1.14	6.79	12.99	13.60	15.98
1e+1a.....	60	0	2.10	0.71	7.68	13.53	17.36	14.09
1a to 1c.....	90	0	0.24	2.72	0.50	1.55	2.33	1.84
1d to 1e.....	90	0	1.90	0.08	6.19	12.53	12.70	13.23
1b to 1d.....	90	0	0.11	4.01	0.19	1.79	4.25	0.17
1e to 1b.....	90	0	1.95	1.88	4.06	6.38	7.75	2.72
1c to 1e.....	90	0	0.48	1.54	5.32	10.57	9.26	15.89
1a to 1d.....	120	0	0.21	2.47	1.14	3.88	1.06	1.07
1b to 1e.....	120	0	0.74	2.95	3.18	5.95	4.00	6.53
1c to 1a.....	120	0	0.65	0.64	5.15	10.70	9.44	13.78
1d to 1b.....	120	0	1.83	1.86	3.79	7.33	6.49	4.26
1e to 1c.....	120	0	0.69	2.04	3.51	5.96	5.62	5.68
1a to 1e.....	150	0	0.84	1.98	3.40	6.82	5.02	6.38
1b to 1f.....	150	0	0.35	1.31	2.95	5.88	3.54	6.47
1c to 1g.....	150	0	0.89	0.09	0.83	3.66	2.82	8.80
1d to 1a.....	150	0	1.72	0.81	1.25	4.66	4.69	7.12
1e to 1b.....	150	0	1.75	0.25	0.00	1.07	1.81	0.99
1f to 1d.....	180	0	0.69	0.77	1.43	0.02	3.10	0.13
1g to 1e.....	180	0	1.41	1.75	0.36	2.27	1.44	3.60
1a to 1f.....	180	0	0.47	0.73	3.16	6.61	4.43	6.34
1b to 1g.....	180	0	1.00	1.22	0.10	1.68	0.36	3.77
1c to 1a.....	180	0	0.94	0.32	1.30	4.64	3.77	8.38
1a to 1g.....	210	0	1.03	0.76	0.60	2.76	1.47	3.95
All stocks...	302	0	0	0	0	0	0	0

TABLE 37

PERCENTAGES OF ERROR FOUND IN STOCK-PRICE INDEX NUMBERS DERIVED FROM SAMPLES OF VARIOUS SIZES

SAMPLE		PER CENT OF ERROR							
Composition	Number of Items	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High	All Dates
1a	30	0	1.84	2.72	6.82	10.79	12.05	4.90	
1b	30	0	2.25	7.08	6.01	9.10	16.14	22.26	
1c	30	0	4.97	2.36	2.48	4.33	1.86	14.09	
1d	30	0	2.02	1.67	4.42	9.93	2.92	8.22	
1e	30	0	3.60	0.60	14.59	14.33	25.86	18.89	
Arithmetic av.	30	.....	2.94	2.89	6.86	9.70	11.77	13.67	7.97
Median	30	.....	2.25	2.36	6.01	9.93	12.05	14.09	7.78
1a+1b	60	0	2.07	2.60	0.14	0.01	3.23	9.81	
1b+1c	60	0	1.27	4.77	1.86	2.53	9.16	4.50	
1c+1d	60	0	1.37	2.01	3.48	7.21	0.59	11.07	
1d+1e	60	0	2.93	1.06	10.29	12.46	16.14	14.38	
1e+1f	60	0	2.93	0.66	11.64	12.98	20.60	12.68	
Arithmetic av.	60	.....	2.11	2.22	5.48	7.04	9.94	10.49	6.21
Median	60	.....	2.07	2.01	3.48	7.21	9.16	11.07	5.83
1a to 1c	90	0	0.33	2.52	0.76	1.49	2.77	1.65	
1d to 1a	90	0	2.65	0.07	9.38	12.02	15.07	11.90	
1b to 1d	90	0	0.15	3.72	0.29	1.72	5.04	0.15	
1e to 1b	90	0	2.72	1.74	6.15	6.12	9.20	2.45	
1c to 1e	90	0	0.67	1.43	8.06	10.14	10.99	14.29	
Arithmetic av.	90	.....	1.30	1.90	4.93	6.30	8.61	6.09	4.85
Median	90	.....	0.67	1.74	6.15	6.12	9.20	2.45	4.39
1a to 1d	120	0	0.29	2.29	1.73	3.73	1.26	0.96	
1b to 1e	120	0	1.03	2.73	4.82	5.71	4.75	5.87	
1c to 1a	120	0	0.91	0.59	7.80	10.27	11.20	12.40	
1d to 1b	120	0	2.55	1.72	5.74	7.03	7.70	3.83	
1e to 1c	120	0	0.96	1.89	5.32	5.72	6.67	5.11	
Arithmetic av.	120	.....	1.15	1.84	5.08	6.49	6.32	5.63	4.42
Median	120	.....	0.96	1.89	5.32	5.72	6.67	5.11	4.28
1a to 1c	150	0	1.17	1.84	5.15	6.54	5.96	5.74	
1b to 1f	150	0	0.49	1.21	4.47	5.64	4.20	5.82	
1c to 1g	150	0	1.24	0.08	1.26	3.51	3.35	7.92	
1d to 1a	150	0	2.40	0.75	1.89	4.47	5.57	6.41	
1e to 1b	150	0	2.44	0.23	0.00	1.03	2.15	0.89	
Arithmetic av.	150	.....	1.30	0.82	2.55	4.24	4.25	5.36	3.09
Median	150	.....	1.24	0.75	1.89	4.47	4.20	5.82	3.06
1f to 1d	180	0	0.96	0 71	2.17	0.02	3.68	0.12	
1g to 1e	180	0	1.97	1.62	0.55	2.18	1.71	3.24	
1a to 1f	180	0	0.66	0.68	4.79	6.34	5.26	5.70	
1b to 1g	180	0	1.40	1.13	0.15	1.61	0.43	3.39	
1c to 1a	180	0	1.31	0.30	1.97	4.45	4.47	7.54	
Arithmetic av.	180	.....	1.26	0.89	1.93	2.92	3.11	4.00	2.35
Median	180	.....	1.31	0.71	1.97	2.18	3.68	3.39	2.21
1a to 1g	210	0	1.44	0.70	0.91	2.65	1.74	3.55	1.83
All stocks	302	0	0	0	0	0	0	0	0

1921, this percentage of error was, however, exceeded in three cases out of five, even though the size of the sample was increased to 120 items.

With the figures in Table 37 at hand, the maker of index numbers of stock prices is enabled to determine what size of sample is required, under market conditions similar to those prevailing during the period 1916 to 1925, to keep an index series within any required limit of error. For example, the figures indicate that, if one is dealing with stock-price figures for dates five years or more from the base and is satisfied with index numbers having errors not greater than 8 per cent, he will, in something like the proportion of four cases out of five, secure the accuracy desired by using a sample of 120 items.

TABLE 38

PERCENTAGE OF ERROR WITH SAMPLES OF DIFFERENT SIZES IN INDICES OF STOCK PRICES CALCULATED FROM AGGREGATE VALUES COMPOSED OF ONE SHARE OF EACH STOCK

Number of Shares in Sample	Average of All Dates	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High
				Per Cent of Error				
30	7.97	0	2.94	2.89	6.86	9.70	11.77	13.67
60	6.21	0	2.11	2.22	5.48	7.04	9.94	10.49
90	4.85	0	1.30	1.90	4.93	6.30	8.61	6.09
120	4.42	0	1.15	1.84	5.08	6.49	6.32	5.63
150	3.09	0	1.30	0.82	2.55	4.24	4.25	5.36
180	2.35	0	1.26	0.89	1.93	2.92	3.11	4.00
210	1.83	0	1.44	0.70	0.91	2.65	1.74	3.55
				Errors Relative to the Thirty-Item Sample Used as a Base				
30	100.00	0	100.00	100.00	100.00	100.00	100.00	100.00
60	77.04	0	71.77	76.82	79.89	72.58	84.45	76.74
90	60.75	0	44.21	65.74	71.87	64.95	73.15	44.55
120	56.44	0	39.11	63.67	74.05	66.91	57.70	41.19
150	38.13	0	44.22	28.37	37.17	43.71	36.11	39.21
180	31.26	0	42.86	30.80	28.13	30.10	26.42	29.29
210	25.76	0	48.98	24.22	13.27	27.32	14.78	25.97

Table 38 presents in compact form the averages found in Table 37, and also shows the results of expressing each error as a relative on the basis of the sample of 30 items. From this table, we see that multiplying the number of items by 6 diminishes the average percentage of error by something more than two-thirds.

Relation of errors to general "law" of error.—The results shown in Table 38 conform but roughly to the general mathematical rule that the per cent of error should vary inversely with the square root of the size of the sample. This lack of conformity is probably due mainly to the fact that the law of error is based upon the behavior of a very large number of items governed by the laws of chance. Under such circumstances, the error would diminish to zero only when the size of the sample reached infinity. In our experiment, however, the totality consists of only 302 items, and the error must disappear when the sample attains this size. This fact of course changes the shape of the curve representing the relationship of the size of the error to the number of items in the sample.

Reducing the error by omitting the low-priced stocks.—Thus far, our investigation of stock-price index numbers has been confined to a study of the behavior of averages of samples selected proportionally from among the entire list of stocks. Is this method of obtaining samples and averaging the items the one most likely to come as near as possible to giving the correct answer to Question 1?

The findings presented in Table 34 indicate that the omission (after arraying the prices per share) of the four extreme items in each sample *increases* instead of *diminishes* the error present in the average. May there not, however, be some other method of attack which will secure index numbers of improved quality?

We know that the method of aggregates emphasizes the importance of those shares which have high prices. We

TABLE 39

THE DEVIATIONS OF DIFFERENT INDEX NUMBERS FROM THE ACTUAL RELATIVES ANSWERING QUESTION 1. SAMPLES SELECTED FROM HIGHER-PRICED STOCKS ONLY

Relatives Computed from	Sample: Number	Sample: Number of Items	Nov. 1916 High	Dec. 1917 Low	Nov. 1919 High	Aug. 1921 Low	Mar. 1923 High	Oct. 1923 Low	Mar. 1925 High	Average of Deviations for 6 Dates
Aggregate of all companies	....	**302**	**100.00**	**71.61**	**107.85**	**66.00**	**104.20**	**84.26**	**111.15**	**0**
Aggregate of samples: Index	2a	31	100.00	86.52	104.83	70.42	102.75	83.64	102.04	
	2b	31	100.00	72.12	109.05	67.11	102.11	83.25	106.17	
	2c	31	100.00	73.44	103.68	70.25	113.14	96.08	126.37	
Deviations	2a	.....	.....	14.91	3.02	4.42	1.45	0.62	9.11	5.59
	2b	.....	.....	0.51	1.20	1.11	2.09	1.01	4.98	1.82
	2c	.....	.....	1.83	4.17	4.25	8.94	1.82	15.22	6.04
	Av. Dev.	.....	.....	5.75	2.80	3.26	4.16	1.15	9.77	4.48
Median items: Index	2a	31	100.00	82.96	96.01	66.36	76.51	64.52	84.80	
	2b	31	100.00	83.98	101.53	74.85	103.87	90.41	99.88	
	2c	31	100.00	73.80	98.85	66.56	102.37	83.68	97.58	
Deviations	2a	.....	.....	11.35	11.84	0.36	27.69	19.74	26.35	16.22
	2b	.....	.....	12.37	3.68	8.85	0.33	6.15	11.27	7.11
	2c	.....	.....	2.19	9.00	0.56	1.83	0.58	13.57	4.62
	Av. Dev.	.....	.....	8.64	8.17	3.26	9.95	8.82	17.06	9.32
Sums of 5 median items: Index	2a	31	100.00	84.73	97.40	65.51	82.55	70.73	87.95	
	2b	31	100.00	84.48	102.29	75.62	101.48	91.83	99.54	
	2c	31	100.00	75.81	97.55	66.53	94.69	83.95	98.13	
Deviations	2a	.....	.....	13.12	10.45	0.46	21.65	13.53	23.20	13.74
	2b	.....	.....	2.87	5.56	9.62	2.72	7.57	11.61	6.66
	2c	.....	.....	4.20	10.30	0.53	9.51	0.31	13.02	6.31
	Av. Dev.	.....	.....	6.73	8.77	3.54	11.29	7.14	15.94	8.90

Median of relatives:										
Index	2a	31	100.00	82.50	96.30	68.30	93.60	77.60	85.40	
	2b	31	100.00	75.80	98.20	73.60	89.80	85.00	92.50	
	2c	31	100.00	71.80	96.10	71.80	95.20	93.40	94.50	
Deviations	2a	.....	.....	10.89	11.55	2.30	10.60	6.66	25.75	11.29
	2b	.....	.....	3.19	9.65	7.60	14.40	0.74	18.65	9.04
	2c	.....	.....	0.19	11.75	5.80	9.00	9.14	16.65	8.76
	Av. Dev.	.....	.....	4.76	10.98	5.23	11.33	5.51	20.35	9.70
Means of relatives:										
Index	2a	31	100.00	89.99	110.13	68.52	97.55	78.27	95.61	
	2b	31	100.00	73.38	117.87	68.70	104.03	84.70	108.85	
	2c	31	100.00	72.52	104.65	70.62	112.33	97.08	125.65	
Deviations	2a	.....	.....	18.38	2.28	2.52	6.65	5.99	15.54	8.56
	2b	.....	.....	1.77	10.02	2.70	0.17	0.44	2.30	2.90
	2c	.....	.....	0.91	3.20	4.62	8.13	12.82	14.50	7.36
	Av. Dev.	.....	.....	7.02	5.17	3.28	4.98	6.42	10.78	6.27
Aggregates of all items except 4 extremes:										
Index	2a	31	100.00	86.75	103.78	66.30	91.55	74.80	89.43	
	2b	31	100.00	74.10	112.44	68.29	98.07	85.51	104.99	
	2c	31	100.00	72.06	99.66	68.79	108.19	92.17	115.13	
Deviations	2a	.....	.....	15.14	4.07	0.30	12.65	9.46	21.72	10.56
	2b	.....	.....	2.49	4.59	2.29	6.13	1.25	6.16	3.82
	2c	.....	.....	0 45	8.19	2.79	3.99	7.91	3.98	4.55
	Av. Dev.	.....	.....	6.03	5.62	1.79	7.59	6.21	10.62	6.31
Geometric averages of original items:										
Index	2a	31	100.00	83.54	99.63	57.89	82.38	59.85	77.62	
	2b	31	100.00	71.34	106.27	63.03	94.40	74.10	84.68	
	2c	31	100.00	70.30	98.82	63.39	95.72	79.96	100.58	
Deviations	2a	.....	.....	11.93	8.22	8.11	21.82	24.41	33.53	18.00
	2b	.....	.....	0.27	1.58	2.97	9.80	10.16	26.47	8.58
	2c	.....	.....	1.31	9.03	2.61	8.48	4.30	10.57	6.05
	Av. Dev.	.....	.....	4.50	6.28	4.56	13.37	12.96	23.52	10.87
Aggregate of all companies:										
Relative	.....	302	100.00	71.61	107.85	66.00	104.20	84.26	111.15	

learned from Table 28 that, in the period under consideration, high-priced stocks rose much less rapidly than low-priced ones. In view of the fact that each low-priced share included eliminates the possibility of taking in a high-priced stock which will have real influence, is it wise to include the cheaper shares in the sample? May it not be better for the purpose at hand to select all sample items from the higher-priced stocks? This plan has been followed in choosing the items included in samples 2a, 2b, and 2c. In so far as possible, the same items were included in 2a, as in 1a, in 2b as in 1b, and in 2c as in 1c. This procedure was followed in order to make the comparison emphasize as much as possible the specific question at issue, namely, the relative influence of high-priced and low-priced stocks. The method followed, then, was to discard from the old sample stocks worth less than $70.00 per share[2] in November, 1916, substituting high-priced stocks in the places of those discarded. The index numbers derived from these new samples of the higher-priced shares appear in Table 39.

A comparison of the total errors arising by this method of sampling with those recorded in Table 34 shows a very distinct superiority on the part of the samples selected from among the higher-priced stocks—in fact, each of the six types or averages have errors well below the corresponding ones appearing in Table 34.

While, then, the evidence indicates that the method of aggregates gives the index series approaching most nearly to the desired results—a finding which accords in part with the commonly held view that the best formula for use with sample data is the one which, when applied to the totality, gives the results sought—our later experiment shows this premise is only partially correct, for we also discover that,

[2] In a very few instances in which a group contained no stock selling at as much as $70.00 in November, 1916, stocks at slightly lower values were included in the sample.

if one desires to secure from a small sample an index number portraying as accurately as possible the fluctuations in the aggregate values of one share of each stock, the lower-priced shares should not be included in the sample. However, the fact should not be overlooked that, as the size of the sample is enlarged, a point must ultimately be reached at which the exclusion of any class or classes of items from the sample will increase the percentage of error in the index series.

COMMODITY PRICE-RELATIVES

The two price-fields most commonly analyzed by means of index numbers are those of securities and of commodities at wholesale. We have just considered ways of answering Question 1—a question concerning stock prices. Let us now turn to Question 2—a query pertaining to the prices of commodities at wholesale.

In the case of the stocks, our totality consisted of price quotations on 302 issues. Similarly, we shall use as a totality for commodity prices the quotations on 336 commodities previously described for the dates

August, 1915
August, 1917
Average for 1919
May, 1920
January, 1922
August, 1925

It will be remembered that all of the dates just listed represent important turning-points in the price average compiled by the United States Bureau of Labor Statistics.

The first problem in this case is exactly the same as in the case of the stock data—namely, to test the accuracy of index numbers computed by the aid of various formulæ from small samples. In the study of wholesale prices, it was decided to use, in every instance, samples consisting of 31 commodities.

With samples of this size, the probabilities are, that the fact that the totality contains only 336 items, instead of being an ideal universe made up of the prices of all commodities sold at wholesale, will affect very little the sizes of the errors, the reason being that the average of the 336 items is not likely to differ materially from the average of the universe. Whenever, however, the samples are increased in size to include a large proportion of the 336 items, there will, of course, be a tendency to minimize the size of the error in those instances in which the formula used in computing the average of the totality is also used in calculating the average of the index.

DEVIATIONS FROM THE SIMPLE ARITHMETIC AVERAGE OF 336 RELATIVES OF COMMODITY PRICES

The first question asked concerning wholesale prices is as follows:

QUESTION 2

What, as compared to 100 as a base, was the relative value at different dates of a basket containing, throughout the period covered, that amount of each of the 336 commodities in the totality, which, at the average price prevailing in 1919, sold for $1.00?

The answer to Question 2 given by the totality.—The exact answer to this query is shown by the items in the first line of Table 32, which are:

August, 1915,	54.62
August, 1917,	92.51
Average for 1919,	100.00
May, 1920,	121.94
January, 1922,	74.28
August, 1925,	81.76

Answers to Question 2 given by various index series.—Tables 40 and 41 show the approximations to this answer given by eighteen proportionally selected samples[3] each averaged in four different ways. The method of calculating the average percentage of error is the same as that used with the stock prices recorded in Table 37. The results are here condensed to save space; as a result, each line in Table 40 represents an average of the percentage errors for five separate dates, while each line in Table 41 shows the result obtained by averaging the percentage errors of 18 separate index numbers.

TABLE 40

AVERAGES FOR FIVE DATES OF THE DEVIATIONS FROM THE SIMPLE ARITHMETIC AVERAGE OF 336 RELATIVES OF WHOLESALE PRICES OF FOUR TYPES OF INDEX NUMBERS EACH DERIVED FROM 31-ITEM PROPORTIONALLY SELECTED SAMPLES

TYPE OR AVERAGE OF RELATIVES	ERROR OF SAMPLE MEASURED IN PER CENT									
	All samples	A	B	C	D	E	F	k	l	m
Simple arithmetic	5.560	3.12	5.52	7.02	19.04	4.09	6.93	3.09	5.89	5.78
Geometric	6.787	4.58	5.19	8.38	2.64	6.22	9.93	7.22	8.48	7.82
Weighted arithmetic	7.642	5.55	7.72	7.45	12.11	6.96	4.39	6.09	6.28	8.01
Median	8.710	7.49	7.30	10.44	6.68	10.29	14.44	6.32	9.23	7.77
		n	o	p	q	r	s	t	u	v
Simple arithmetic	5.560	1.60	5.74	3.04	1.45	3.22	2.84	7.85	4.89	8.95
Geometric	6.787	4.32	6.72	6.42	4.80	6.94	5.93	12.75	6.23	7.60
Weighted arithmetic	7.642	1.37	10.97	10.14	8.12	7.59	6.99	11.29	5.49	11.03
Median	8.710	5.44	9.32	7.57	4.97	10.02	9.11	13.52	9.45	7.42

Errors of index numbers compared for stock prices and commodity prices.—The figures show that the errors in the index numbers of wholesale prices do not differ materially in size from the errors in the index numbers of stock prices;

[3] The method used in selecting these samples is described in Chapter V.

TABLE 41

AVERAGE OF THE PERCENTAGES BY WHICH 18 INDEX NUMBERS DEVIATE FROM THE SIMPLE ARITHMETIC AVERAGE OF 336 RELATIVES OF WHOLESALE PRICES. THE INDEX NUMBERS HAVE BEEN DERIVED BY VARIOUS FORMULÆ FROM EACH OF FOUR SETS OF PROPORTIONALLY SELECTED SAMPLES CONTAINING 31 COMMODITIES EACH

TYPE OR AVERAGE OF RELATIVES	NUMBER OF ERRORS AVERAGED	AVERAGE PER CENT OF ERROR IN						
		Average for Five Dates	Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
Simple arith...	18	**5.56**	11.23	6.35	0.00	3.53	3.41	3.27
Geometric	18	**6.79**	12.16	8.34	0.00	4.34	4.42	4.67
Weighted arith...	18	**7.64**	10.78	6.75	0.00	9.19	6.13	5.35
Median...	18	**8.71**	13.41	13.73	0.00	5.41	4.80	6.20

although it must be admitted that the behavior of the stock prices is somewhat more erratic. Occasionally, in each case, the price movement in a single item is sufficiently radical to introduce a large error into the average of the sample. For example, in the case of stock prices, the mean of the relatives of Sample 1b was more than 244 per cent in error in March, 1925, while the simple arithmetic average of Sample D of commodity prices at wholesale was 81 per cent away from the correct location in August, 1915. Such instances are, of course, exceptional, but they prove how undependable the behavior of an arithmetic average index number based upon a limited number of commodities may be. Such indices, when derived from small-sized samples, are, then, far from qualifying for the title "instruments of precision."

Comparative errors of different index numbers.—We see from Table 40 that, when the errors for the five dates are averaged, the simple arithmetic average of relatives is more

accurate than the geometric in the case of 15 out of the 18 samples. The geometric, in turn, is better than the weighted arithmetic in 11 cases out of 18. The weighted arithmetic is likewise better than the median in 11 cases out of 18. The most decisive indication of superiority is that given by the first comparison. These results indicate, that, for the purpose of answering Question 2, the simple arithmetic average of relatives completely outclasses the other three formulæ.

Reference to the figures in Table 41—figures which show for each separate date the results of averaging the errors of 18 index numbers—confirms the superiority of the simple arithmetic average for the purpose at hand, for, in all five periods, it gives greater accuracy than the geometric average; in four out of five periods, it gives a smaller error than the weighted arithmetic average; and, in every case, it shows distinct superiority over the median.

Relation of size of error to time-distance from base.—The figures in Table 41 indicate that the tendency for errors to increase in size in proportion to the time-distance from the base is much less marked in the case of commodity prices than in the case of stock prices. True, the errors are, in general, much larger in August, 1915, than in August, 1917, but those in August, 1925, average little, if any, larger than those in May, 1920. This is another illustration of the fact that prices of stocks and commodities differ in their behavior.[4]

Again, just as in the case of the question pertaining to stock prices, the evidence indicates that the best formula to use with sample data is the one which gives the correct answer when applied to the totality. As previously stated, this conclusion accords with the assumption tacitly made by

[4] The reader may be interested in comparing the tendencies shown by this index with those of the index of the dispersion of commodity price relatives. For the latter, see pages 256 to 286 of the report of the National Bureau of Economic Research prepared by Dr. Frederick C. Mills and entitled *The Behavior of Prices.*

most authorities on index numbers. It will be remembered, however, that, in the case of stock prices, it was found necessary to modify this conclusion to some extent. Perhaps we shall also find modifications necessary in the case of commodity prices.

Relation of size of error to size of sample.—We know that, when the formula applied in calculating the index number is identical with the correct formula for the totality, the error must approach zero as a limit as the size of the sample is increased. A test has been made of the relationship between the size of the sample and the size of the error arising when the simple arithmetic average of sample relatives is used to answer Question 2.

In order to extend the curve as far as possible, recourse was had to random sampling—the procedure being as follows: All the cards were shuffled and 9 samples of 31 items each were chosen by purely chance methods. The samples were combined by using the same process illustrated in Tables 35, 36, and 37. The figures presented in Table 42 are, in each case, the averages of the percentages of error of ten separate samples. By this process of averaging numerous samples, a rather high degree of regularity has been attained—a regularity which justifies the feeling that the curve represents the normal tendency in this particular type of data.

We find that, just as was true in the case of stock prices, the errors diminish regularly as the sample grows larger. Since, however, index numbers derived from small samples have smaller errors in the case of commodity than in the case of stock prices, the slope of the latter curve is necessarily steeper than that of the former.

The figures in Table 42 show that, in the present attempt to answer Question 2, the error in a small sample of 31 items averages about 7 per cent, but may occasionally run to double that figure. To reduce the average error to 2 per cent

requires a sample containing over 200 items, and, even with a sample of this size, the error may run as high as 5 per cent in some cases. The figures in the table are helpful in establishing the size of sample necessary to secure any desired degree of accuracy in the index, but the fact should be kept in mind that, if the sample were selected from a very large universe instead of from a totality containing only 336 items, the percentage of error would not be artificially forced to zero at the 336 item limit, and hence the errors would not diminish so rapidly with an increase in the size of the sample.

TABLE 42

THE AVERAGE PERCENTAGE DEVIATIONS FROM THE SIMPLE ARITHMETIC AVERAGE OF 336 RELATIVES OF GROUPS OF TEN INDEX NUMBERS OBTAINED BY AVERAGING RANDOM SAMPLES VARYING IN SIZE

NUMBER OF ITEMS IN EACH SAMPLE	AVERAGE PER CENT OF ERROR SHOWN BY TEN SAMPLES						
	Average for Five Dates	Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925
31	**6.95**	13.10	6.52	0.00	2.40	7.98	4.75
62	**5.49**	11.66	4.47	0.00	1.57	5.73	4.03
93	**4.52**	9.89	3.06	0.00	1.32	4.99	3.34
124	**3.74**	8.42	2.42	0.00	1.15	3.92	2.79
155	**3.16**	7.12	2.18	0.00	1.04	3.12	2.35
186	**2.65**	5.85	2.15	0.00	0.86	2.52	1.86
217	**2.10**	4.65	1.71	0.00	0.73	2.05	1.37
248	**1.60**	3.45	1.50	0.00	0.68	1.38	0.98
279	**1.10**	2.06	1.23	0.00	0.68	0.89	0.66
310	**0.70**	0.71	1.22	0.00	0.67	0.30	0.60
336	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00

Thus far our study of the merits of various kinds of index numbers has been directed to ascertaining which best con-

form to the SIMPLE arithmetic average of all the relatives in the totality of 336 items. We shall now turn to a consideration of a second question also pertaining to the relative prices of commodities at wholesale.

DEVIATIONS FROM THE WEIGHTED ARITHMETIC AVERAGE OF 336 RELATIVES OF COMMODITY PRICES

Different commodities are not produced in equal quantities or consumed in equal quantities, hence we usually consider that they differ in importance. As a result, it is presumably true that the demand for an answer to Question 2 is not so great as the demand for an answer to the question which follows, namely, Question 3. This question is:

QUESTION 3

If it is assumed that, for every commodity, a dollar's worth in 1919 constitutes one unit of quantity, and, if it is further assumed that a huge basket be filled with 336 commodities, the number of units of each being proportional to the percentage which the value of that commodity constituted in 1919 of the total value of products in the Continental United States (as estimated by the United States Bureau of Labor Statistics, from Census and other data) for how much, relative to the value in 1919, would the contents of this basket sell in each of the five months listed?

Methods of computing the answer.—Two different procedures may be followed in arriving at the answer to the above question; first, the method of aggregates; second, the arithmetic average of relatives, using as weights the values in 1919. Both, for reasons thoroughly explained in earlier chapters, give identical results.

In the particular study now being described, it was found simpler to begin the computation by calculating relatives on

the base 1919. In the construction of weighted averages of relatives, the weights already prepared by the United States Bureau of Labor Statistics were used. Each price relative was multiplied by the per cent which that Bureau estimates each commodity to have constituted of the total value of all products in the year 1919.

The answer to Question 3 given by the totality.—The addition of the 336 products for each year, and the reduction of the totals thus obtained to relatives, using the base 1919 as 100, gives the exact answer to the question asked —an answer which is shown in the first line of numbers in Table 33. We know, then, that, under the conditions laid down in the question, the relative values of identical baskets of commodities at the selected dates were as follows:

August, 1915,	50.75
August, 1917,	91.86
Average for 1919,	100.00
May, 1920,	123.96
January, 1922,	73.11
August, 1925,	80.11

Methods used in constructing index numbers.—In making this test, the same eighteen proportional samples used in answering Question 2 (namely, A, B, C, D, E, F, k, l, m, n, o, p, q, r, s, t, u, and v), the selection of which was described in the last two chapters, were used in the initial experiments.

Four different averages or types were computed for each sample. The deviation from the weighted arithmetic average of the 336 commodities was computed in the case of each average of each sample for each date. Following the same plan used in preparing the data for Table 37, these deviations were next converted to percentages by dividing them by the corresponding relatives of the totality. So many individual deviations were thus obtained that it seems best to present only the average percentage deviations of the

eighteen index numbers obtained from the eighteen samples. These average errors are shown in Tables 43 and 44.

TABLE 43

AVERAGES FOR FIVE DATES OF THE DEVIATIONS FROM THE WEIGHTED ARITHMETIC AVERAGE OF 336 RELATIVES OF INDEX NUMBERS OF WHOLESALE PRICES DERIVED BY FOUR DIFFERENT FORMULÆ FROM PROPORTIONALLY SELECTED SAMPLES, EACH CONTAINING 31 COMMODITIES

TYPE OR AVERAGE OF RELATIVES	ERROR OF SAMPLE MEASURED IN PER CENT									
	All samples	A	B	C	D	E	F	k	l	m
Simple arithmetic...	**5.39**	2.85	4.43	5.13	22.34	5.60	5.39	4.63	4.19	4.70
Geometric.....	**6.47**	3.50	3.60	6.58	5.47	4.31	8.19	5.36	6.68	6.02
Weighted arithmetic.	**6.95**	4.08	6.51	5.87	14.58	5.68	3.34	5.83	5.29	8.06
Median.....	**7.36**	7.14	5.88	8.61	5.23	8.41	12.76	6.93	7.44	5.85
		n	o	p	q	r	s	t	u	v
Simple arithmetic...	**5.39**	3.25	5.35	2.49	3.54	1.48	3.91	5.46	4.43	7.88
Geometric.....	**5.47**	3.96	5.34	4.60	2.91	5.10	4.77	11.06	4.98	6.05
Weighted arithmetic.	**6.95**	2.17	9.11	7.57	6.67	6.44	8.40	9.85	5.07	10.62
Median.....	**7.36**	5.19	8.15	5.86	4.27	8.24	7.24	11.20	8.22	5.87

TABLE 44

PERCENTAGE DEVIATIONS FROM THE WEIGHTED ARITHMETIC AVERAGE OF 336 RELATIVES OF 18 INDEX NUMBERS OF WHOLESALE PRICES DERIVED BY DIFFERENT FORMULÆ FROM EACH OF 4 SETS OF PROPORTIONALLY SELECTED SAMPLES CONTAINING 31 COMMODITIES EACH

TYPE OR AVERAGE OF RELATIVES	NUMBER OF ERRORS AVERAGED	PER CENT OF ERROR IN						
		Average for Five Dates	Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug 1925
Simple arith.	18	**5.39**	9.87	6.17	0	3.46	4.00	3.46
Geometric	18	**5.47**	6.98	7.78	0	5.54	3.57	3.50
Weighted arith...	18	**6.95**	7.60	6.72	0	9.29	6.05	5.09
Median..	18	**7.36**	7.25	13.61	0	6.23	4.67	5.05

Errors little affected by time-distance from base.—The figures in Table 44 tend to strengthen the view that the percentage errors of indices of the prices of commodities at wholesale have no tendency to increase in magnitude as the indices depart in point of time from the base date.

Errors of different index series compared.—While the fact just mentioned is of interest, relationships of much greater moment are revealed by the figures presented in Tables 43 and 44, which show the comparative accuracy, on the average, of index numbers computed respectively by the four different formulæ used in the experiment. We see that, surprisingly enough, it is still the oft-condemned *simple* arithmetic average of relatives which yields the index numbers having the smallest errors, even though the comparison is now with the *weighted* arithmetic average of the totality. Furthermore, this difference in favor of the simple average cannot be ascribed to chance, for Table 43 shows that, in twelve out of the eighteen samples, the simple arithmetic average shows a percentage of error smaller than that given by the geometric average and, in 15 out of 18 samples, it is superior to the weighted arithmetic average. Furthermore, a study of Table 44 reveals the fact that, at three dates out of the five covered, the simple arithmetic averages showed errors smaller than those of the geometric averages; while, in four months out of five, the simple arithmetic averages gave results more nearly accurate than the weighted arithmetic averages.

Why weighting increases the error of the arithmetic average of a small sample.—How can these surprising phenomena be explained? Clearly it would be impossible for the index numbers computed by calculating the simple arithmetic averages of relatives to show the smallest deviations from the two series of relatives of totalities answering respectively Questions 2 and 3, were it not for the fact that the series for both these totalities follow approximately the

same course. As a matter of fact, the two sets of relatives lie very close together at all dates, as the following figures show:

Relative of Totality	August 1915	August 1917	Average for 1919	May 1920	January 1922	August 1925
Simple.....	54.62	92.51	100.00	121.94	74.28	81.76
Weighted.....	50.75	91.86	100.00	123.96	73.11	80.11

At every date except August, 1915, the divergence is small. Presumably, in normal times, when price fluctuations are not so marked as in the period covered, the courses of the relatives of the two totalities would diverge even less. We need not fear, then, that we are basing our conclusions upon a chance relationship between these two totalities. There is, on the contrary, every reason to believe that, as a rule, the simple arithmetic average of the relatives comprising a small sample measures more accurately than does the *weighted* arithmetic average of the sample, the course of the *weighted* arithmetic average of all the relatives in the totality.

But how does this remarkable state of affairs come about?

Since the unweighted average gives greater importance than the weighted average to the unimportant items, one is naturally led to inquire whether the explanation is not found in a tendency of the prices of the more important commodities to behave in a fashion more erratic than that characterizing the course of the less important commodities. However, this hypothesis is quickly overthrown, for a study of Table 33 shows that the deviation of the group average from the course of the totality was distinctly less at every date in the case of important than of unimportant articles. It appears, then, that the erratic behavior of the weighted averages cannot be due to the comparative characteristics of

important and unimportant commodities included in this particular totality, but must, instead, have its origin in the fact that weighting is employed. The weakness of the weighted arithmetic average seems to arise from the fact that the price of nearly every commodity has a tendency to vary widely and that any marked deviation of a single important item from the trend is emphasized by the weighting, and so affects the average of the sample unduly. To illustrate this principle, let us suppose that a single item is weighted so heavily that it dominates the sample. In that case, since the deviations of the other 30 items tend to cancel each other, the movements of the one heavily-weighted commodity will be left to determine the course of the index. Since the variability of any one item is likely to be much greater than the variability of an average of 31 items, the index governed by the one commodity is likely to fluctuate rather widely, and the chances are, of course, that its movements will not correspond at all closely with those of the totality.

Comparative effects of individual weighting and group weighting.—In computing the weighted arithmetic averages of commodity price relatives with which we have just been dealing, the weights used have, in every instance, been the figures assigned by the United States Bureau of Labor Statistics as representing the percentages which the values of the outputs of the respective commodities constituted of the total value of all commodities in the base year 1919. The question arises as to whether this method of weighting is not responsible for the relatively large size of the errors found in the weighted arithmetic averages of the sample relatives. The fact has been brought out in an earlier chapter that, as a general rule, weight should be assigned, not in proportion to the importance of the sample item, but, instead, in proportion to the importance of the group which it represents. It seemed reasonable, then, to suppose that, if the

sample items had been weighted according to the importance of the groups in which they happened to be included, the errors of the index numbers might have been materially reduced. To test this hypothesis, it was decided to compute new index numbers from six of the same samples used in the studies heretofore recorded. The method used in computing these new indices is as follows:

1. Each original price relative (on the base 1919) was multiplied by the weight assigned to it by the United States Bureau of Labor Statistics as representing its relative importance in 1919.
2. The products thus obtained for all sample items selected from any one of the 15 groups listed in Table 33 were summated. This process was repeated for each date.
3. The sum of the products at each given date was multiplied by the ratio of the weight for the *group* in 1919 (as recorded in Table 33) to the sum of the weights of sample products in 1919.
4. The final products arrived at in Step 3 were summated for each of the dates covered, and the sums thus obtained were reduced to relatives on the base 1919.

The index numbers obtained by this system of group weighting appear in Table 45.

Strangely enough, the results prove to be just opposite to those which the rule calls for. In only 8 out of 30 cases are the errors as small when group weighting is used as when each commodity is weighted according to its own importance. What is the explanation of this anomaly?

The answer seems to be that, despite the fact that the classification was made on a basis which appeared to be thoroughly logical, the prices of commodities in any one group have not actually been dominated by any common price-making force, hence there has been little homogeneity in the price movements of the articles in the group. In respect to price, then, any individual commodity is no more representative of the group than of the totality. Under such circumstances, the commodity cannot be considered as repre-

sentative of the group into which it happens to fall, and there is, therefore, no point in applying the group-weight to its price; indeed, to do so is likely to make matters worse instead of better.

TABLE 45

COMPARATIVE PERCENTAGES OF ERROR IN INDEX NUMBERS RESULTING RESPECTIVELY FROM WEIGHTING RELATIVES ACCORDING TO THE IMPORTANCE OF THE INDIVIDUAL COMMODITY AND THE IMPORTANCE OF THE GROUP CONTAINING THE COMMODITY

SAMPLE	TYPE OF WEIGHTING	PER CENT DEVIATION FROM THE WEIGHTED ARITHMETIC AVERAGE OF ALL 336 RELATIVES						
		Aug. 1915	Aug. 1917	Average for 1919	May 1920	Jan. 1922	Aug. 1925	Average for five dates
A	Group	12.30	2.83	0.00	2.71	2.93	1.10	**4.37**
	Individual	12.97	2.05	0.00	2.53	2.37	0.49	**4.08**
B	Group	0.34	8.52	0.00	16.41	1.31	1.52	**5.62**
	Individual	0.45	5.70	0.00	21.13	2.39	2.87	**6.51**
C	Group	4.67	9.10	0.00	3.27	7.34	32.70	**11.42**
	Individual	4.85	7.12	0.00	0.63	2.30	14.46	**5.87**
D	Group	36.83	16.39	0.00	5.75	10.18	10.41	**15.91**
	Individual	31.74	17.77	0.00	5.62	9.08	8.66	**14.58**
E	Group	1.46	10.33	0.00	12.97	4.09	9.92	**7.75**
	Individual	5.04	6.00	0.00	12.88	3.65	0.85	**5.68**
F	Group	6.25	9.06	0.00	6.07	2.56	7.28	**6.24**
	Individual	0.79	5.44	0.00	3.61	0.72	6.13	**3.34**
Average of errors of six samples	Group	10.31	9.37	0.00	7.86	4.74	10.49	**8.55**
	Individual	9.31	7.35	0.00	7.73	3.42	5.58	**6.68**

If, for example, sugar happens to be the only sample item from the group entitled "Foodstuffs of Greatest Importance," it is given a weight equivalent to 14.9 per cent of

the total for the sample.[5] Whenever the price of sugar happens to fluctuate violently, this heavy weight will cause it to have great influence upon the weighted average of the sample, often introducing a large error into the index. The only effect, then, upon an index of wholesale prices, of substituting group weighting for individual weighting, is to emphasize erratic price fluctuations and hence to increase the errors of the index numbers. This happens because the prices of individual commodities do not tend to be representative of the groups in which they have been placed. Were they representative, the effect would probably be the reverse of that shown by the present experiment.

Arithmetic averages often vitiated by erratic movements of one or two price-relatives.—The evidence obtained by actual experiment shows quite conclusively that, in practice, with a small sample, the *unweighted* rather than the weighted arithmetic average of relatives gives the closest approach to the *weighted* arithmetic average of the totality, while the median of the relatives makes a showing distinctly worse than that of the weighted arithmetic average of the relatives. However, the use of the median in the case of a small sample ought not to be condemned too hastily, for it must be remembered that an index number computed by averaging the items in a small sample is always in danger of being thrown far from the true path by the action of one or two wildly fluctuating items; and this source of error can be avoided by using the median. For example, in one instance, the error in the index computed by taking the simple arithmetic average of the sample climbs to about 81 per cent. In no case, however, does the index calculated by aid of the median contain an error as large as 21 per cent. The question arises as to whether it is not possible to devise a formula which will give an index series having an error approaching the low limit set by the unweighted arithmetic average, while at the same time retaining the

[5] See Table 33.

virtues of the median in protecting the average from serious distortion by the erratic behavior of one or two items. This question will be discussed on a later page.

Relation of size of error to size of sample.—As just stated, the figures in Tables 43 and 44 prove that the simple arithmetic average of relatives is the average which, when used with a small sample, comes nearest to answering Question 3 correctly. Since, however, this formula, when applied to the totality, does not give the correct answer, it is obviously true, that, as the size of the sample is increased, a point must eventually be reached when the weighted average of relatives becomes superior. At just what size of sample does this happen? Table 46 helps to answer this question.

The figures in this table compare the average percentage deviations from the weighted arithmetic average of relatives of two sets of index numbers, both obtained from ten random samples, each consisting of 31 items. The two sets are:

1. Simple arithmetic averages of relatives.
2. Weighted arithmetic averages of relatives.

When the percentages for the five dates are averaged, and the averages thus arrived at are compared, we see that the first class to show the lesser error for the weighted arithmetic average is that containing 124 items. However, the normal crossing point of the two curves apparently is at a point considerably higher than 124, for, at four out of five dates, the simple arithmetic average is, for the particular purpose at hand, superior, even when the sample contains as many as 186 items. The size of sample, then, with which it is best to use the unweighted arithmetic average of relatives in attempting to answer Question 3, is indicated by this experiment to be one containing not more than 200 items, and possibly not more than 125 items. For larger samples, the weighted arithmetic average is distinctly better.

We have now noted the effect that increasing the number

TABLE 46

AVERAGE PERCENTAGES OF DEVIATION FROM THE WEIGHTED ARITHMETIC AVERAGE OF ALL RELATIVES OF THE WEIGHTED AND UNWEIGHTED AVERAGES OF SAMPLES OF VARIOUS SIZES CHOSEN AT RANDOM

NUMBER OF ITEMS IN SAMPLE	NUMBER OF SAMPLES AVERAGED	AVERAGE OF PERCENTAGES OF ERROR					
		Average for five dates	Aug. 1915	Aug. 1917	May 1920	Jan. 1922	Aug. 1925
		Simple Arithmetic Average					
31	10	**6.67**	11.17	6.75	2.05	8.11	5.26
62	10	**5.34**	9.48	4.65	1.55	6.39	4.65
93	10	**4.62**	8.89	3.35	1.38	5.39	4.08
124	10	**4.14**	8.69	2.64	1.25	4.52	3.59
155	10	**3.78**	8.59	2.61	1.18	3.46	3.07
186	10	**3.49**	8.44	2.58	1.05	2.73	2.66
217	10	**3.29**	8.40	2.15	1.02	2.22	2.65
248	10	**3.25**	8.54	2.07	0.97	2.01	2.65
279	10	**3.17**	8.40	1.94	0.96	1.90	2.65
310	1	**3.17**	8.39	1.94	0.97	1.90	2.67
336	1	**2.73**	7.63	0.71	1.63	1.60	2.06
		Weighted Arithmetic Average					
31	10	**8.15**	7.63	7.94	6.89	11.49	6.79
62	10	**5.80**	5.90	5.69	3.69	7.74	5.99
93	10	**4.81**	5.03	5.07	3.10	5.77	5.07
124	10	**3.79**	3.50	4.19	2.52	4.37	4.38
155	10	**3.03**	2.43	3.35	2.16	3.43	3.78
186	10	**2.54**	2.29	2.82	1.50	3.05	3.06
217	10	**2.07**	2.07	1.91	1.41	2.63	2.32
248	10	**1.50**	1.45	1.42	1.08	1.97	1.55
279	10	**1.06**	0.96	1.28	0.95	1.33	0.80
310	1	**0.43**	0.32	1.00	0.70	0.08	0.06
336	1	**0.00**	0.00	0.00	0.00	0.00	0.00

of items in the sample has in reducing the size of the error in three separate cases. The results are summarized in Table 47.

TABLE 47

COMPARATIVE EFFECTS OF INCREASING THE SIZE OF THE SAMPLE UPON THE AVERAGE ERRORS FOUND IN THE ANSWERS TO VARIOUS QUESTIONS

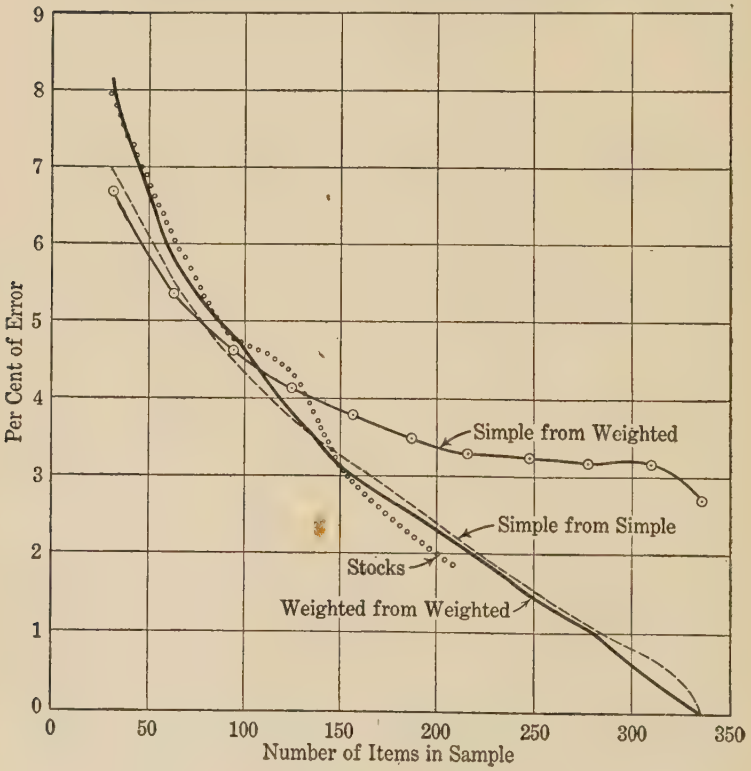
NUMBER OF ITEMS IN SAMPLE	DEVIATION FROM THE			
	Aggregate of 302 stock-prices *of the* aggregate of the sample stock-prices	Simple arithmetic average of 336 relatives of commodity prices *of the* simple arithmetic average of the sample relatives	Weighted arithmetic average of 336 relatives of commodity prices *of the* average of the sample relatives	
			Simple	Weighted
30 or 31	7.97	6.95	6.67	8.15
60 or 62	6.21	5.49	5.34	5.80
90 or 93	4.85	4.52	4.62	4.81
120 or 124	4.42	3.74	4.14	3.79
150 or 155	3.09	3.16	3.78	3.03
180 or 186	2.35	2.65	3.49	2.54
210 or 217	1.83	2.10	3.29	2.07
248	.....	1.60	3.25	1.50
279	.....	1.10	3.17	1.06
310	.....	0.70	3.17	0.43
336	.....	0.00	2.73	0.00

This comparison shows that, for samples of 150 items or less, the errors in the four cases are strikingly similar, both in size and slope. However, as the size of the sample approaches that of the totality, the deviation of the simple arithmetic average of the relatives in the sample from the weighted arithmetic average of the relatives in the totality,

fails to diminish in size as do the others. The reason, of course, is that, since the formula used in computing the index number is different from that applicable to the total-

CHART II

RELATION OF THE AVERAGE ERROR OF INDEX NUMBERS TO THE SIZE OF THE SAMPLE FROM WHICH THEY WERE COMPUTED



ity, the error is not approaching zero at the point where the sample increases to 336 items.

It is undoubtedly true that, when the same formula is

used for both the sample and the totality, the size of the error in the larger samples is affected by the fact that the error is artificially forced to zero at 336 items—the assumed universe of items. It is impossible from the facts at hand to say how the shape of the curve recording the errors would actually be affected by making the totality many times as large. The appearance of the curve itself, however, indicates that its location would probably be changed little, if at all, in the case of samples containing less than 125 items, and that, even with samples made up of as many as 200 items, the location of the curve is not much different from what it would be if the universe were infinitely large.

Comparative merits of random and proportional sampling.—In the studies just described, two methods of sampling have been used—the random and the proportional. The latter method of sampling is much more complicated and laborious. Does it give results sufficiently better in quality to justify the expenditure of the additional effort involved? To answer this question, Table 48 has been constructed. In it appear averages of the errors obtained by the two methods from groups of samples.

TABLE 48

COMPARATIVE ERRORS RESULTING RESPECTIVELY FROM RANDOM AND PROPORTIONAL SAMPLING

Index Numbers Used in Each Case Are Simple Arithmetic Averages of 31 Relatives

AVERAGE OF THE TOTALITY FROM WHICH THE DEVIATIONS ARE CALCULATED	METHOD OF SAMPLING	NUMBER OF ERRORS AVERAGED	AVERAGE PER CENT OF ERROR					
			Average of Five Dates	Aug. 1915	Aug. 1917	May 1920	Jan. 1922	Aug. 1925
Simple arithmetic (Question 2).....	Random	10	**6.95**	13.10	6.52	2.40	7.98	4.75
	Proportional	18	**5.56**	11.23	6.35	3.52	3.41	3.27
Weighted arithmetic (Question 3).....	Random	10	**8.15**	7.63	7.94	6.89	11.49	6.79
	Proportional	18	**5.39**	9.87	6.17	3.46	4.00	3.46

We see that, in answering Question 2, the proportional method of sampling gives results which, when compared to those obtained by random sampling, are found to be better in four out of the five cases. The same proportion holds in answering Question 3. In both instances, the average error is lower for the proportional sample. Clearly, then, when a small sample is used, it is worth while to take the trouble necessary to see that the sample items are selected on a proportional basis.

Improving the sample by eliminating erratic items.—The experimental work done thus far seems to prove conclusively that, in computing index numbers from small samples with a view to answering Question 3, it is better to use the simple arithmetic average of the relatives than it is to depend upon the weighted arithmetic average, the geometric average, or the median. Nevertheless, the fact remains that the arithmetic average, while usually satisfactory, is occasionally made entirely non-representative by the presence of one or two wildly-fluctuating items. Both the median and the geometric average are less subject to the influence of such exceptional variations. There is, then, much to be said in favor of using them, even though they show larger average errors.

Is it not possible, however, to devise some method which will combine the advantages of the median and the arithmetic average?

The solution suggesting itself is to discard the extremes of the sample and then obtain the arithmetic average of the central group of items. However, the experiments along this line with stock prices indicated that results are not satisfactory when the central group of items is too small. Why not, then, after having arrayed the cards, simply discard a few of those at the extremes of the array, and proceed to average the remaining price-relatives?

This plan of attack was tried with samples A, B, C, D, E,

and F. The results showed that, to discard *any* given number of items from the extremes of the sample *increased* the error.

The suggestion was then made by Dr. Maurice Leven that better results might be obtained by discarding only those sample items showing the highest proportional deviations from the central tendency. Experiments were conducted in accordance with this suggestion. In making these tests, it was decided to use the median of the sample as the point from which to calculate the deviations and to take the inter-quartile range—that is, the distance between the quartiles [6] —as the measure of dispersion. Tests were made with various degrees of dispersion from the median. It was finally discovered that, in dealing with samples of commodity prices at wholesale, the nearest approach to answering Question 3 was secured when all price-relatives deviating from the median by more than six times the inter-quartile range were discarded from the sample before averaging the relatives. The marked diminution in the average of the errors in the six samples of 31 items each is shown by the figures in Table 49.

The best index number for answering Question 3.—To sum up, we may say that the results of our inquiry show that, of all the index numbers tested, the best one to use with a small sample in an endeavor to answer Question 3 —that is the one which comes nearest to paralleling the course of the weighted arithmetic average of all of the relatives in the totality—is the simple arithmetic average of relatives of a sample selected proportionally, and with all items eliminated therefrom which deviate from the median by an amount greater than six times the inter-quartile range.

[6] When several, or many, items are arrayed, the quartiles are the items one-fourth, two-fourths, and three-fourths of the way through the array, beginning with the lowest item. The median is identical with the second quartile. The difference between the first quartile and the third quartile is called the inter-quartile range.

TABLE 49

AVERAGE OF THE PERCENTAGE DEVIATIONS AT FIVE DIFFERENT DATES OF THE SIMPLE ARITHMETIC AVERAGE OF SAMPLE RELATIVES FROM THE WEIGHTED ARITHMETIC AVERAGE OF 336 RELATIVES.

SAMPLE	ERRATIC ITEMS ELIMINATED		AVERAGE ERROR OF SAMPLES CONTAINING 31 ITEMS (PER CENT)
	Number of items eliminated at any date	Average error of reduced samples (per cent)	
All samples	7	4.20	7.63
A	1*	3.26	2.85
B	1†	4.11	4.42
C	0	5.13	5.13
D	3‡	4.40	22.33
E	2§	2.89	5.68
F	0	5.39	5.39

* In the sample for August, 1917.
† In the sample for May, 1920.
‡ Two in the sample for August, 1915, and 1 in August, 1917.
§ One in the sample for August, 1915, and 1 in August, 1917.

If a larger sample is to be used, the items should be selected in the same way, and the same erratic items should be eliminated. In this case, however, it seems wiser to use the weighted arithmetic average of the relatives in the sample. The elimination of the extreme variants among the price-relatives would, of course, lower throughout its course the curve showing the percentage of error to be expected with samples of various sizes.

SUMMARY

1. This chapter is devoted to a record of experiments made to test the merits of various formulæ used in the computation of index numbers.

2. The merit of an index series is indicated by the close-

ness with which its course conforms to that of a series of relatives representing the course of the totality.

3. It is usually assumed that the formula required to answer a given question from complete data is the best one to use with sample data. This is not necessarily true.

4. The first series of experiments was made with index numbers designed to approximate a series of relatives representing the course of the aggregate prices of 302 shares of an equal number of stocks. The chief facts discovered in this connection are as follows:

(a) Errors in index numbers of stock prices tend to increase as the time-distance from the base-date increases.

(b) The larger the sample used in constructing the index number, the smaller is the error.

(c) With a small sample, the most accurate index numbers are obtained by including in the sample a disproportionately large share of the higher-priced stocks, computing the aggregates of the share-prices, and reducing these aggregations to relatives.

5. The second set of experiments was made with a view to finding which formula yields the index series coming nearest to equalling at various dates the simple arithmetic averages of 336 commodity price-relatives. The following generalizations are indicated by the results of the experiments:

(a) In occasional instances, the errors of indices of stock prices are much greater than the errors of indices of commodity prices. On the average, however, the errors are about equal in size. With either kind of prices, small samples are likely to yield index numbers having very large errors.

(b) In general, of the averages tested, the simple arithmetic average of relatives gives the index series hav-

ing the smallest errors. The geometric average of relatives is, however, less subject to the chance of large errors. The same may be said of the median of relatives.

(c) The percentage of error in indices of commodity prices shows little tendency to grow larger as the time-distance from the base-date increases.

(d) Just as in the case of stock prices, the percentage of error grows less as the size of the sample increases.

6. The third set of experiments was designed to ascertain what formula gives the index series conforming most closely to a series of relatives representing the *weighted* arithmetic averages of 336 relatives of commodity prices. The following principles were established:

(a) With samples containing less than 125 items, the best results obtained from any of the tests were secured by using the simple arithmetic average of relatives as the formula for computing the index number. With samples of 200 items, or over, the weighted arithmetic average of relatives proved superior. With samples of 125 to 200 items there appears to be little choice between the two methods.

(b) The reason that, on the average, the weighted arithmetic average of relatives gives poor results with small samples is that, occasionally, an erratic price movement occurs in an important commodity, and a single instance of this kind causes a large error in the entire sample.

(c) With the grouping used in this experiment, substituting the weight of the group from which sample items are selected in place of the combined individual weights of the samples in the group increases the error of the index number.

7. The curves representing the relationship between the size of the error and the size of the sample are much the same for commodity prices as for stock prices.

8. Proportional sampling gives index numbers having smaller errors than those derived from random sampling.

9. The best way of protecting index numbers from the disastrous effects produced by a few erratic price-changes, is to eliminate from the sample, before averaging, all items showing a variation from the median of more than six times the inter-quartile range.

QUESTIONS AND PROBLEMS

1. Suppose that fixed-base relative prices are available for all the stocks on the market and that these relatives are averaged to form an index series; frame the exact question that would be answered

(a) By the harmonic average.
(b) By the geometric average.
(c) By the median.
(d) By the simple arithmetic average.
(e) By the weighted arithmetic average.

Which of the questions correspond with the queries commonly made concerning stock prices?

2. Are the formulæ proposed for use in the construction of index numbers few or many? Is it easy or difficult to design a new formula? Try it for yourself.

3. Is it always feasible to test the merit of an index series by comparing its course with that of the series of relatives representing the totality?

4. Does one always have the option of increasing the size of the sample until the error is reduced enough to bring it within the required limits?

5. Is the probability that the correct formula for the totality is the best formula to apply to the sample data greater when the sample is large or when it is small?

6. Why does the error of an index of stock prices grow decidedly larger as its time-distance from the base increases, while the index of commodity prices does not show a similar tendency to any marked extent?

7. According to the law of chance errors, how great a fractional reduction in the size of the error would be brought about by in-

creasing the size of the sample from 30 to 120 items? How closely is this rule borne out by our experiments with indices of

(a) Stock prices?
(b) Commodity prices?

8. If there were close correlation between the importance of commodities and the respective movements of their prices, would you expect the simple arithmetic average of the sample to have an error larger or smaller than that of the weighted arithmetic average?

9. Why is weighting likely to increase the degree of error in the average of a small sample?

10. If the error of an index series can be reduced by eliminating from the sample extremely erratic items, why can it not be still further reduced by eliminating all erratic items?

11. Does the fact that, with a small sample, the simple arithmetic average of relatives gives the closest approach to the weighted arithmetic average of all relatives in the totality prove that it is always the best formula to use if only a small sample is available?

BIBLIOGRAPHICAL NOTE

For a study of the theory of error as applied to averages of sample data, see the references at the end of Chapter IV and also the following:

Arthur L. Bowley, *Elements of Statistics* (1907), pp. 305–315.

F. Y. Edgeworth, *Journal of the Royal Statistical Society*, July, 1923, pp. 570–590; September, 1923, pp. 343–351; July, 1925, pp. 557–575.

CHAPTER VIII

MEASURING THE PURCHASING POWER OF MONEY

To measure the purchasing power of money is sometimes assumed to be the only important use of index numbers.—Perhaps the most frequent use of index numbers in the past has been in connection with attempts to measure changes in the *general* purchasing power of money or of gold. Correa M. Walsh says on page 116 of his book *The Problem of Estimation*:[1] "In averaging price-variations, the purpose or object is given: it is to measure variations in the exchange-value or purchasing-power of money. . . . To get at this true variation of the exchange-value of money is the sole object or purpose of averaging the price variations. . . . There is no other object or purpose—that is, no other main or final purpose. The single object or purpose has been stated. It is specific. No further varietal subdivision of it is needed, nor is it possible, so far as the method of averaging the price variations is considered, including the method of weighting them."

According to the principles expressed on earlier pages of the present report, Walsh's view is decidedly erroneous, but it well illustrates the dominance of the *general* price index concept in the minds of students of this field.

While, in his book on *The Making of Index Numbers,* Professor Irving Fisher enumerates many uses to which index numbers are now being put, he states that: "The origi-

[1] P. S. King and Son, London, 1921.

nal purpose of index numbers—to measure the purchasing power of money—will remain a principal, if not the principal, use of index numbers." [2] His treatment of the problem throughout his book indicates that, while he was engaged in its preparation, this particular use of index numbers remained uppermost in his mind from start to finish.

There is every reason to believe that this circumscribed concept of the use of index numbers has been largely responsible for the errors into which many inquirers have fallen. It is this particular use of index numbers about which most discussion has centered,[3] though a number of writers have, from time to time, emphatically called attention to the necessity of having a different index number for each problem to be solved.

Although the measurement of the general purchasing power of money is only one of the thousands of possible ends to be served by index numbers, it must be admitted that it is an extremely important end. The use of index numbers for this particular purpose is, indeed, of such consequence that it seems worth while to devote this chapter to its consideration.

The chief uses of an index of the general purchasing power of money.—How do economists hope to employ an index of the general purchasing power of money after it is computed? It appears that, in the main, they are actuated by one or more of the following four reasons:

1. They desire to ascertain the effect upon the price-level of variations in the factors affecting the supply of money; *i.e.*, they wish to see, for example, what, if any, results have been produced by changes in the amount of
   (a) Gold in the world.
   (b) Gold in the Treasury, or in bank reserves.
   (c) Other coin or money in the Treasury, or in bank reserves.
   (d) Gold in circulation.

[2] See pages 368-369 of Fisher's book.

[3] As an indication of this, see the discussions referred to in F. Y. Edgeworth's *Papers Relating to Political Economy*, Chapters H to O.

(e) Coin and paper money in circulation.
(f) Bank deposits subject to check.
(g) Bonds and other credit instruments outstanding.

2. They wish to note the effect upon the price-level of changes in the factors affecting the demand for money, as, for example, variations in the volume of trade or business activity.
3. They are interested in arriving at a series of relatives or index numbers which may be used as a standard for deferred payments which will prove satisfactory even in cases in which payment is due many decades after the date of the contract.
4. They hope to obtain a series of relatives or index numbers which can be widely used for the purpose of "deflating" figures representing amounts of money as stated in dollars current at the given date.

There is no doubt that the general purchasing power, or, in other words, the value of money, like the value of any other commodity, is determined by the forces of supply and demand. Demand for money is, of course, affected by all the uses to which money is put.

Chief uses of money.—Among the purposes for which money is often employed, the following may be mentioned:

1. Coins are used as ornaments.
2. Certain rare coins are preserved as souvenirs.
3. Money is hoarded for future needs.
4. The average person keeps a rather definite amount of money constantly on hand for current daily expenses.
5. Money is collected by banks, and frequently by individuals, in order to have it available to loan to others.
6. Money is kept by banks to pay loans coming due, and to meet the calls of depositors.
7. Money is collected by individuals in order to pay off loans that are approaching maturity.
8. Money is required by banks and merchants in order to make change.
9. Money is gathered together in order to buy specific commodities.

Use for purchase of commodities is one most studied.—Although nine different uses are enumerated above, all but the first two are directly or indirectly related to the particular function of money to which economists have devoted most attention in recent years, that is, to the purchase of

commodities. It appears, indeed, that, in some cases, economists have assumed either that all other calls for money, such as, for example, the need for money to make change, to lend to others, or to pay off debts, are of rather minor importance, or else that these requirements either vary but little from time to time, or vary in proportion to the changes in the demand for money as a medium of exchange.[4] Since the payment of loans requires large sums of money, and since a loan is frequently paid off at a time far distant from the date when the proceeds of the loan were used for the purchase of goods, this assumption may, in many cases, do considerable violence to the facts. The effect upon the purchasing power of money of these miscellaneous uses is a subject well worthy of investigation.

As just stated, scientific inquiry thus far has been limited primarily to studying the use of money as a medium for buying goods. Much impetus was given to investigations in this field by Professor Irving Fisher's book entitled *The Purchasing Power of Money* and especially by the device which he uses therein to measure the different factors involved. This device is the equation of exchange.

The equation of exchange.—In its simplest form, this equation is written

$$MV = PT$$

the terms being defined as follows:

M equals the average number of dollars (including coin, paper money, and bank deposits subjects to check), circulating in the given territory during the period in question.

V equals the number of times, on the average, each of

[4] Professor Morris A. Copeland's studies in the *Proceedings of the American Statistical Association*, March, 1929, p. 146, and in the *Journal* of the same organization, June, 1929, p. 109, constitute exceptions to this rule.

these dollars was paid out during the period for the purchase of goods.

T equals the number of units of commodities exchanged for money during the period.

P equals the average price paid per unit for all commodities bought for money during the period.

It will be observed that the equation is merely a mathematical form of saying that the total price paid for all commodities bought for money is equal to the total value of those commodities. There can be no dispute about this fact.

Two sides of equation of exchange refer to different dates.—Unfortunately, one encounters a serious difficulty whenever one attempts to substitute for the terms of the equation figures representing actual transactions. The right-hand side of the equation, "PT", is a statement of the total value of all goods sold, while the left-hand side of the equation, "MV", evidently represents the total of the money used in the process of paying for the commodities. As business is actually transacted, the goods sold at a given date are not paid for until a considerably later period, and it is only when the time of payment arrives, that the stress impinges upon the money supply. Evidently, then, in any given field, the two sides of the equation refer to two different dates, the time-distance between which varies according to the nature of the commodities dealt in, and the customary method of payment in each case. To know, then, what actual rate to use in the equation of exchange in order to make $MV = PT$, it is evidently necessary to know first the weighted average length of time intervening between the date of sale and the date of payment. As far as the present writer is aware, no research work along this line has yet been undertaken, and hence the extent of the lag is unknown.

The total value of goods sold.—Were it not for this time-difference, the substitution of actual data for the terms of

the equation would, at first thought, seem to be extremely easy. Evidently "PT" equals the total value of goods (including services) sold in a given period. There is no need to distinguish between wheat, iron, milk, clothing, bonds, stocks, wages, salaries, rents, or fees. The total value of everything sold is included. If, in the interval under consideration, any commodity is sold eight times, its value must appear eight times. Whenever, then, the necessary records are available, the measurement of "PT" presents no unusual statistical obstacles.

The physical volume of goods sold.—The difficulty begins when we attempt to separate "P" from "T". To do this, it is obviously necessary to find the total number of units of all commodities sold during the given period. These units must be defined in physical terms. Clearly, to add together the numbers sold of barrels of flour, years of labor, and shares of stock, is so contrary to the most elementary principles of arithmetic as to be out of the question. Yet, if an average price is to be obtained, some method of combining flour, labor, and stocks must be devised.

Some makers of index numbers seem to think they have avoided the difficulty by choosing one period as the base, reducing all prices to relatives on that base, and thereafter dealing with these relatives instead of with the original data. This device accomplishes nothing except to obscure the real nature of the operations involved. If the relatives are averaged without weighting, it is equivalent to assigning to each commodity equal total value in the base year, and assuming that the physical quantities in other years bore to each other the same proportional relationships prevailing in the base year. Since, in practice, both the quantities and the values of the respective commodities sold differ greatly in the base year as well as in other years, there is no defensible reason for giving them identical weights.

If relatives are computed, and these relatives are weighted

in a certain manner, the process becomes equivalent to the use of aggregates—a fact already demonstrated repeatedly. The direct employment of aggregates rather than the indirect use of weighted relatives makes the procedure simpler, and easier to understand and explain. Before aggregates can be computed, it is, however, necessary to reduce all the items to terms which can be added. How can this best be done?

The best quantity units for use in measuring the general purchasing power of money.—Clearly, no physical unit is of any value for statistical purposes unless it remains unchanged in size and nature throughout the period studied. Can units of this type be devised for individual commodities which units are also of such a nature that those for different commodities can be added together to secure an aggregate physical quantity for all commodities?

In his book on *The Purchasing Power of Money,* Professor Fisher presents a simple method of arriving at units having both these qualities—a method which is decidedly the best yet devised for the purpose. He selects a given period as a base and then defines as one unit the physical quantity of any commodity sold in that period for $1. This unit remains exactly the same at all dates.

To illustrate the principle, let us imagine a community in which all dealings have been for cash and only three commodities have been traded in—namely, flour, labor, and stocks in the Consolidated corporation. Transactions during the last two years were as follows:

Commodity	Unit	1926			1927		
		Price per unit	Number sold	Total value	Price per unit	Number sold	Total value
Flour	Barrel	$ 8	180	$ 1,440	$ 6	200	$ 1,200
Labor	Year	1,000	14	14,000	1,100	16	17,600
Consolidated stock	Share	50	60	3,000	40	160	6,400
Total				$18,440			$25,200

In the example here given, if the year 1926 is chosen as the base, a unit of flour consisted of $\frac{1}{8}$ barrel, a unit of labor $\frac{1}{1000}$ of a year's work, and a unit of stock $\frac{1}{50}$ of a share. On this basis, the total sales may be pictured as follows:

Commodity	Unit (Dollar's Worth in 1926)	1926			1927		
		Price per unit	Number of units	Total value	Price per unit	Number of units	Total value
Flour	$\frac{1}{8}$ bbl.	$1.00	1,440	$ 1,440	$ 0.75	1,600	$ 1,200
Labor	$\frac{1}{1000}$ yr.	1.00	14,000	14,000	1.10	16,000	17,600
Consolidated stock	$\frac{1}{50}$ share	1.00	3,000	3,000	0.80	8,000	6,400
Total			18,440	$18,440		25,600	$25,200
Average		$1.00			$0.985		

By calling every amount selling for $1.00 in 1926 one unit, we find that the $18,440 spent in that year bought 18,440 units at a cost of $1.00 per unit while, in 1927, the $25,200 spent, paid for 25,600 units, making the average price $0.985. On this basis, then, one can say either that the price level fell 1.5 per cent or that the dollar would buy $\frac{1}{0.985}$ or 101.5 per cent as many goods in 1927 as in 1926.

Which base is best is a question impossible to decide.— But, in choosing a year for a base, there is no particular reason for preferring 1926 to 1927. Will the answer be the same if the latter year is taken as a base? It will not, for the physical units of the various commodities are all changed in size when a new base date is adopted. If we use 1927 as the base year, we find that one dollar would buy $\frac{1}{6}$ barrel of flour, $\frac{1}{1,100}$ year of labor, and $\frac{1}{40}$ share of stock; hence these become the new units of quantity and the computation is as follows:

Commodity	Unit (Dollar's Worth in 1927)	1926			1927		
		Price per unit	Number of units	Total value	Price per unit	Number of units	Total value
Flour	¼ bbl.	$ 1.33	1,080	$ 1,440	$1.00	1,200	$ 1,200
Labor	$\frac{1}{1000}$ yr.	0.91	15,400	14,000	1.00	17,600	17,600
Consolidated stock	$\frac{1}{40}$ share	1.25	2,400	3,000	1.00	6,400	6,400
Total	.....	.....	18,880	$18,440	.....	25,200	$25,200
Average	.....	$0.977	.....	.....	$1.00		

We see that the shift of base to 1927 causes an increase in the number of units recorded for 1926 and a decrease in the number recorded for 1927. As a result, we now find that the average price per unit has increased from $0.977 to $1.00 or 2.4 per cent instead of undergoing the fall of 1.5 per cent shown when 1926 was used as the base year. Which answer is correct? How can one answer the query: "Did prices in general actually rise or fall between 1926 and 1927?"

The reply must be that it is impossible to give a defensible answer to the question as stated because the question itself is vague and indefinite. To make it definite, it is imperative to include in the question a statement of the base with which all comparisons are to be made.

The simplest procedure is to follow the plan just illustrated of choosing arbitrarily some one date as a base and adhering to it thereafter. Some authorities believe, however, that, since any given date is likely to be non-representative of conditions during a long period, more meaningful comparisons may be obtained by using as a unit of any given commodity an average of the amounts purchasable for one dollar at the beginning and at the end of the interval covered. It must be admitted that, by so doing, one is likely to make the weighting more typical of the entire period. However, this end is still better served by averaging the amounts sold for a dollar at all the recorded dates during

the period covered and using the average as a unit. To do this, however, materially adds to the work involved in the computation.

Furthermore, the fact should not be overlooked that any such averaging process also increases the difficulty of explaining the meaning of the resulting figures—in other words, the question answered by the figures is made more complex and harder to understand. When the dollar's worth in the base year is taken as a unit, the question may be stated thus: "If a dollar's worth of any commodity in the base year is assumed to constitute one unit what, in other years, was the average price per unit paid for all units sold?" If, however, the suggested averaging process is adopted, the question must be changed to read thus: "If the amounts of any given commodity sold for a dollar in the various periods quoted are averaged, and if the average thus obtained is defined as one unit of the given commodity, what, in the various years covered, was the average price per unit, when all commodities are taken together?"

From the facts thus far stated, we may, then, draw the following conclusions concerning the choice of a base and the units to be employed:

1. In order to arrive at the number of units sold, it is necessary to know the quantity of each article constituting a unit.
2. The most convenient unit appears to be the amount worth one dollar in the base year.
3. Presumably, with the varied list of commodities necessary for a satisfactory index of this type, it makes little difference which date is chosen as a base—at any rate there is no way of deciding whether, for this purpose, one date is better or worse than another.
4. The procedure frequently suggested of averaging the quantities for several years and using the averages as units complicates the interpretation of the index, and it is doubtful that it makes the figures obtained more reliable for use in calculating the movement of the general price level by aid of the equation of exchange.

Measurement of physical volume of trade difficult because many important commodities are not standardized.—Even after the choice of base and the method of defining the unit have been satisfactorily determined, it is by no means easy to put the definition into practice. True, one can readily enough ascertain the weight of a given grade of wheat which could be bought for a dollar at a specified time. Unfortunately, however, a large proportion of the commodities bought and sold are not, like wheat, classified into well-defined grades, the specifications of which remain unchanged over long periods of time. It is, for example, much less simple either to define the unit or to describe the physical characteristics of that fraction of a suit of clothes which could be purchased for a dollar at a given date and to obtain, for a later time, the price of the same fraction of a suit of the same quality, Similar difficulties appear when one attempts to value real estate—and real estate constitutes the largest single item of national wealth and is widely traded in.

The difficulties are multiplied when one takes securities into account. At first thought, it may seem very simple to deal with such definite things as shares of stocks, but, unfortunately, as Professor Wesley C. Mitchell has pointed out,[5] the significance of a share of stock is always changing. Even if no new stock is issued, the nature of the physical property which underlies and gives value to the stock is constantly shifting. Generally, in addition, the problem is complicated by issues of new stock and various other forms of re-financing.

If one is interested merely in measuring short-time cyclical swings or violent changes in the value of "P", such as those occurring between 1916 and 1921, the influence of any variations in the nature of the standard of commodities such as housing, stocks, etc., may be so small as to be inconse-

[5] See the *Journal of Political Economy*, July, 1916, p. 625.

quential in comparison with the enormous variations brought about by other forces. If, however, one desires to measure "P" with such a degree of accuracy that it can be used satisfactorily as a long-time standard of deferred payments, it is necessary to eliminate from the items used in making the calculation, all those not subject to reasonably accurate standardization.

General purchasing power of money must perforce be estimated from sample data.—In either case, such a large proportion of sales are of unstandardized commodities that it is clearly impossible to estimate the quantity "PT" from anything like complete data. The best that can be done is to utilize sample data and to infer that, on the average, the prices of non-standardized commodities vary in about the same proportions as do more or less comparable prices of standardized commodities.

Sample should be widely inclusive.—Owing to the limitations of the data available, it appears that, to make a really satisfactory index, it is necessary to include practically all items of consequence for which it is possible to secure, throughout the period which it is desired to cover, both the total value of the product and the number of units sold. It is, of course, imperative that the comparison be made in every instance upon the basis of units practically identical in nature.

Best to weight sample data.—Many authorities have held that, if changes in the value of "P" are to be measured by a study of sample data, commodities of trivial value make just as good samples for this purpose as do the most important commodities, the argument being that there is no more reason to believe that important commodities respond more accurately or readily than do unimportant commodities to the forces affecting the purchasing power of money. On the other hand, it is contended with sound reason that, if one commodity, so important that it dominated the entire market,

rose but little in price during the period in which the currency was being inflated, nearly all of the demand arising from the presence of the excess currency would be available to lift the price of the relatively unimportant commodities, and that, therefore, they would vary much more widely than the dominant commodity and hence much more widely than the average of all commodities. While this logic appears to be sound, the fact probably is that, in a country like the United States, there are no commodities which, by any stretch of the imagination, can be considered dominant; hence there is some reason for arguing that one commodity constitutes as good a sample as another, in so far as measuring the effect of various factors on the purchasing power of money is concerned. However, as we have seen, with large samples the error is reduced by using the same weighting with the sample that is applicable to the totality. In attempting the measurement of the general purchasing power of money, it seems desirable to use a large sample. Under these circumstances, weighting is presumably distinctly advantageous.

Formula to be used in computing a general price index.—With complete data at hand, all that is required to measure "P" in the equation of exchange is to divide the total value of goods sold by the total number of physical units sold. Because of the complex nature of the data involved, it does not seem feasible to do any experimental work in determining the best formula to use in computing an index number of the general price level. One is therefore forced to depend upon deductive reasoning for a solution of the problem. The data involved are so heterogeneous in their nature that it would be dangerous to rely upon a small sample—in fact, the only feasible way to arrive at a dependable index number seems to be to include as large a part of the total volume of trade as possible.

We know that the larger the percentage of all existing data used in the construction of an index number, the

greater is the probability that the formula applying to the totality is the correct one to use with the sample data. It appears highly probable, then, that, for use with a *large and comprehensive* sample, the formula illustrated in Table 50 is the one best adapted to the purpose at hand, since, when applied to complete data, it has the advantage of being reasonably simple in its meaning, something which cannot be said of certain more complex formulæ which have been seriously proposed as substitutes.

As to what formula applied to a *small* sample of prices would give the index series which would come nearest to conforming in its movements to those obtained from complete data by the formula described above, it is impossible to say. This question could be answered only after lengthy experimentation, conducted along the lines illustrated in the last chapter.

Variable weights are desirable.—Since our conclusion is that, with a comprehensive sample, the best procedure, in computing an index number of the general price-level suitable for use in the equation of exchange, is to follow the same method which would be used were the data complete, and since the commodities traded in are constantly changing as regards the number of units of each sold, it is apparently desirable to use varying weights in the construction of the index of the general purchasing power of money. Strangely enough, those who have devoted most attention to the problem seem commonly to have overlooked this fact.

The use of different physical quantities at different dates makes the formula for measuring changes in the general purchasing power of money differ radically from any of those which we have previously studied, for, in answering the queries heretofore considered, it has, in every instance, been found imperative to keep constant, at all times, the physical quantity of each of the commodities under consideration.

Substitution of new commodities for old.—The maker of any index covering a long period of time will discover that customs of consumption and therefore of production, change from time to time, with the result that commodities important at one date become of little consequence a few years later. At present, for example, the automobile has largely taken the place of the carriage; the electric light has replaced gas; and the radio has been substituted for the phonograph. Any index number that is intended to give a relatively complete picture of price changes during a long period of time must take into account and provide for important substitutions of this type. To accomplish this, some form of chain system appears to be imperative. What is probably the best method of splicing is illustrated by the following hypothetical example, the assumption made therein being that, in the years covered, quotations are available only for the commodities entered in the table. This particular variety of pieced index will probably come nearer than any other to portraying the actual movements of the general price level.

Two separate indices of purchasing power desirable—trend index and cycle index.—Since it is desirable to have the sample used in measuring the purchasing power of money as representative as possible, and since certain items like corporate shares are suitable for use in short-time comparisons but are unfitted for inclusion in indices designed to measure long-time trends, it appears that the best way to secure maximum accuracy is to compute one index to show the general trend and another to picture the cyclical and short-time fluctuations. The sample used in computing the index designed to give the correct long-time trend should, however, be based upon a set of price data differing materially in its composition from the set of price data used in constructing the index of short-time price fluctuations. Before entering into the make-up of each index, it is necessary

TABLE 50

CORRECT PROCEDURE FOR COMPUTING AN INDEX OF THE GENERAL PRICE LEVEL BY CHAINING RELATIVES

Method: Spliced Relatives of Aggregates Using Changing Quantities

Commodity	1925			1926		
	Value of total sales	Number of units sold	Average price per unit	Value of total sales	Number of units sold	Average price per unit
A	$1,804	304	.....	$1,750	260	
B	650	50	.....	720	60	
C	918	102	.....	968	88	
D	240	60	.....	210	80	
Total	$3,612	516	$7.00	$3,648	488	$7.70
Relative, base 1925	.....	1.00	.....	.....	1.10	
Chain index, base 1925	.....	.....	1.00	.....	.....	1.10
	1926			1927		
A	$1,750	260	.....	$1,536	240	
B	720	60	.....	900	76	
D	210	80	.....	220	90	
E	280	40	.....	454	60	
F	1,216	82	.....	1,070	84	
Total	$4,176	522	$8.00	$4,180	550	$7.60
Relative, base 1926	.....	.....	1.00	.....	.....	0.95
Chain index, base 1925	.....	.....	.....	.....	.....	1.04

to consider the peculiarities of each of certain important groups of commodities.

The rent of houses should appear in the cyclical index only. —A considerable proportion of the national income is spent for the rent or purchase of houses, hence, it would seem especially desirable to include the price of housing in an index of the general price level. Both the United States

Bureau of Labor Statistics and the National Industrial Conference Board have demonstrated the feasibility of securing quotations which are reasonably satisfactory when used for the purpose of measuring month-to-month changes. It seems, then, to be both desirable and feasible to give the price of housing a place in the list of commodities entering into the short-time index of the general price-level.

However, the quality of housing varies slowly but surely as time passes. It is extremely difficult, therefore, to secure a quotation on the rent of a five-room house to-day and be at all certain that the house is in any way comparable with the five-room structure for which the rent was quoted thirty years ago. Plumbing, heating, lighting, and interior finish are all likely to have changed radically. It is not impossible that the cost per cubic foot may have doubled.

Under such circumstances, it seems best, in constructing the long-time index, to be satisfied to use as substitutes quotations on building materials and, if possible, piece-rates for the wages of certain types of building laborers, if necessary, weighting these items somewhat more heavily than would otherwise be the case in order to make up for the omission of house rents from the list of quotations.

The price of labor.—It is far from easy to ascertain what changes have occurred in the price of a unit of labor. What constitutes a unit of labor? Shall we use the rate per week, the rate per day, the rate per hour, or the rate per piece of work done? The first three of these rates will have a trend differing from that of the prices of commodities whenever the efficiency of labor as a producing agent is changing. To secure homogeneity in the general trend index, it appears necessary, then, in so far as labor engaged in so-called productive industries is concerned, to confine wage quotations to piece-rates. Such wage figures, unfortunately, cannot be secured in quantities sufficient to be very helpful in constructing a general price-index. Records of

wage rates per hour, day, or week are, however, available for many industries. In the case of *direct* services such as those performed by house servants, barbers, physicians, and dentists, the rate per hour appears to be suitable for inclusion in a long-time price index, the ground for such inclusion being that the hour of service is a fairly standard commodity which does not change much with the passage of time.

The argument against including hourly wage rates in such fields as manufacturing, mining, and transportation, rests largely upon a utilitarian rather than upon a theoretical basis. If the general price index is to constitute a satisfactory standard of deferred payments, it is desirable that no unnecessary heterogeneity be introduced into the data. Now piece-rate wages, and the prices of the commodities produced by the labor, tend to vary in unison without regard to changes in productivity per man-hour. Such is far from being the case with hourly rates. It appears, then, to be easier to understand a long-time trend of the purchasing power of money if the only wage rates of *indirect* [6] labor entering into it are piece-rates.

Since, however, the productivity of labor changes but slowly, hourly or weekly wage-rates for *indirect* labor (which are often much easier to obtain than piece-rates) are entirely satisfactory for use in measuring cyclical and other short-time fluctuations of the general price level.

Ought bond prices to be considered in measuring the general price level?—The question of what to do with bond prices is a rather puzzling one. Bonds are bought and sold just like other commodities. They come in standard units. They are regularly quoted. Their purchase or sale certainly affects materially the supply of money available for purchasing other commodities.

On the other hand, the price of a bond depends primarily

[6] By indirect labor, we mean that which is not employed in directly satisfying human wants.

upon interest rates and the probability that the bond will be paid when due, and hence is affected but little by changes in the supply of circulating medium—the force which is always important and often dominant in the case of practically all other prices.

If the purpose in computing the index of the general price level is to facilitate the measurement of the size of the various quantities entering into the equation of exchange, the likelihood that the results will be accurate is materially increased by including bond prices among the data. If, on the other hand, the effort is to arrive at a standard of deferred payments, it is just as evident that bond prices should be excluded from consideration.

Whatever the course that the maker of general price-index numbers decides to follow, it is clear that he should be careful to explain in connection with the index series, whenever it is published, whether bond prices have or have not been used in its construction.

Prices of stocks ought to be included in the cyclical index only.—Prices of corporate shares are not suited for inclusion in the index showing the long-time trend, since, as previously shown, a share does not represent the same thing at widely separated dates, and hence any index of the purchasing power of money based partly upon the prices of common stock is likely to have an erroneous slope imparted to its trend.

However, for short spaces of time, prices of common shares, if corrected for obvious changes due to stock-dividends and the like, are perfectly satisfactory for inclusion in the data entering into a general price-index. An especial reason for the inclusion of prices of common shares lies in the fact that it is impossible to standardize and hence to secure satisfactory price-comparisons in point of time covering lands, mines, buildings, machines, power-plants, embankments, railways, and similar objects, and these objects make

up the bulk of the wealth of corporations and a considerable part of the total wealth of the nation.

In view of the facts just stated, it seems then, that it is highly desirable to take the prices of common stocks into consideration whenever one attempts to measure the cyclical fluctuations and other short-time movements of the general price-level. Since changes in the money supply affect prices of common stocks in about the same way as they affect commodity prices in general, the objections registered against the inclusion of bond prices do not apply.

Preferred stocks occupy a position half-way between that of stocks and bonds. Since the limitations on their dividend rates usually prevent the shares from rising in price indefinitely, it seems best to class preferred stocks with bonds, for the purpose at hand.

Feasible to compute a satisfactory general price-index.—Because of the fact that great groups of important commodities are unstandardized, and because different persons are certain to disagree as to just which classes of more or less standardized articles are to be utilized in making up the index, it appears certain that no index of the general price level can be computed which will be satisfactory to everyone. This is far from saying, however, that it is useless to attempt measurements in this field. It is probably easier to secure an index which throws light on short-time swings than to obtain one which reveals the actual trend; but even the latter may be approximated successfully.

Commodities to be included in each type of index.—From the discussion pertaining to the various types of commodities, it appears that the general price-index designed to picture the long-time trend should be constructed from a comprehensive sample of prices from the following fields:

1. Time wage-rates of persons rendering direct service.
2. Piece wage-rates for laborers engaged in indirect labor—that is, labor producing goods or rendering services not directly satisfying human wants.

3. Prices of tangible goods sold at wholesale.
4. Prices of tangible goods sold at retail.
5. Prices of well-defined services, such as railway transportation, electric power, etc.
6. Bonds should be included only in those cases in which the object is to measure the different quantities entering into the equation of exchange.

In calculating the index designed to cover short-time fluctuations only, the following types of items should be added:

7. Time wage-rates of labor engaged in *indirect* processes.
8. House rents.
9. Prices of common stocks.

Combining the cycle index with the trend index.—For reasons stated in connection with the discussion of various commodity groups, it appears that a better picture of changes in the general purchasing power of money can be secured by computing separately a trend index and a cycle index, and then combining the two, than could be arrived at by any other method. The necessary procedure may be outlined as follows:

The deviations of the index of short-time fluctuations from the trend of that index itself should first be computed. These deviations ought then to be added to the *trend* of the index series designed to measure the long-time movements of prices. The composite result should give the closest possible approach to the end sought, and should yield an index series portraying correctly, not only the major fluctuations in the general price level, but also approximating its true trend.

The best standard of deferred payments.—Recently, economists have been giving serious consideration to the problem of stabilizing the value of the dollar. This can only be done on the basis of some kind of an index number, hence it has been necessary to consider what type of index number is fitted best for this particular purpose. It appears that the main object in stabilizing the purchasing power of

the dollar is to guarantee that deferred payments, when made, shall be in dollars always having the same value. The most desirable index for use as a guide for stabilizing the dollar is, then, the same as that price index which is best from the standpoint of deferred payments.

Dr. Maurice Leven of the New York State Housing Commission makes the novel proposal that, for purposes of stabilization, an index of average wages per hour would give the nearest approach to justice for all. The reasons for his view may be summarized thus:

At present, in the United States, we are, fortunately, in an era of increasing production per capita; in other words, with the passage of time, an hour's labor comes to produce a larger amount of commodities than was formerly the case. With such conditions prevailing, it is only just that the creditor should participate in the economic progress of society, and he should, therefore, when his loan is repaid, receive a larger amount of commodities than he loaned. This arrangement works no injustice to the debtor, for, under conditions of increased productivity, it is easier for him to obtain a given number of dollars than it was when he borhowed the money. Since, when productivity increases, hourly wage-rates rise, an index of average wages per hour would serve as a satisfactory criterion of the economic progress of the nation.

Dr. Leven believes that there is still another advantage in using an index of hourly wages as a basis for fixing the value of the dollar. With the dollar stabilized along these lines, wage contracts covering long periods of time might be made between employers and labor unions, for, manifestly, without any adjustments in such contracts, the laborers would share in the general progress of society as a whole; whereas, if wage contracts were fixed on a commodity basis, it would be necessary to modify them constantly in order to take account of the changing productivity of labor.

Although it must be admitted that there are certain marked advantages in following Dr. Leven's suggestion and using as a standard an index of the price of labor per hour, there are also serious objections to the adoption of this plan. One is that his proposal is contrary to the prevailing conception of justice between debtors and creditors, for most people probably believe that, in addition to his interest, the creditor is entitled to receive only the same amount of commodities which he loaned. A more serious objection is probably one that is political rather than economic in nature, namely, that, with prices stabilized on this basis, it would be difficult to convince the average trade unionist that labor was really participating in the economic improvement of society; for the fact that wages, as measured in cents per hour, would remain unchanged year after year, would probably give the impression that labor was not advancing, but was being held down for the benefit of other classes.

Because of the difficulties just mentioned, it appears wisest, in choosing a standard to be used as a basis of stabilization, to conform to the view most commonly held—namely that the standard should be such as to make the amount of goods repaid equal to the amount loaned. However, even if we accept this principle as a basis of action, we find that, before it can be applied, many details must be worked out. Most contracts are made in terms of money and not in terms of goods. Different creditors sold different goods in order to obtain the money which they loaned. No one can tell what type of goods the various creditors will desire to purchase years hence when the respective loans come due, and there is no way of knowing what kinds of goods the individual debtors would need to dispose of in order to obtain the money to pay their respective debts. Under such circumstances, we may be quite sure that the index number giving absolute justice to any one creditor will probably not be exactly fair either to other creditors or to the man who owes

him money. Similarly, the index which proves just, as regards one debtor, may be distinctly unfair to another debtor. The best, therefore, that can be done is to fix upon some general standard which, so far as can be foreseen at present, is likely to give the maximum of satisfaction to the largest number of people. What ought the characteristics of such a standard to be? To make the question more specific, what articles ought to be included in the list used in the construction of the index numbers? Since, as has just been pointed out, it is entirely impossible to give exact justice to individual debtors and creditors, the best that can be done is to adopt a standard which will be as nearly equitable as may be for these classes as a whole. From this standpoint, it appears that the logical policy is to make the index as inclusive as possible. The most inclusive index is that which measures the general purchasing power of the dollar —in other words, the average price of all goods bought and sold in the United States—and it is this index which, therefore, seems best adapted to the purpose at hand.

If, however, this general index number is to be used as a standard of deferred payments, it is entirely illogical to include the price of bonds among those prices entering into such an index, for, even though bonds are bought and sold like other commodities, and therefore call for money, the fact remains that they are, after all, nothing more or less than promises to pay *money* at some future date. The factors affecting bond prices are, therefore, radically different from those affecting the prices of other commodities. There is no reason, for example, why even the most radical inflation of the currency should affect the price of bonds. Now, the fundamental reason for establishing a standard of deferred payments is to make certain that the debtor pays, and the creditor receives, a fixed amount in goods—not a fixed amount of money. Manifestly, then, it is not sound practice to include bond-prices among the data used in computing any

index designed for use in stabilizing the currency, or as a standard for deferred payments.

The same objections do not hold in the case of stocks, for stock-prices, like those of most other commodities, are affected by the supply of circulating medium. However, as we have seen, stocks are not well suited for inclusion in an index number intended for comparisons over a long period of time, the reason being that they are non-standardized articles, a share in one year not representing the same thing as a share at another date.

Index of price of direct goods probably best standard of deferred payments.—If in the construction of the long-time standard of deferred payments we omit from the basic data quotations on stocks and bonds, the question arises as to whether it may not also be desirable to eliminate from the list the prices of all indirect goods. Manifestly, indirect goods are not, as a rule, wanted for themselves; instead they constitute merely a means to an end. The ultimate demand is not for pig-iron, wheat, cotton, or lumber, but is instead for automobiles, flour, dresses, and houses—things that satisfy one's wants directly. As many economists have pointed out, the values of indirect goods are merely more or less imperfect representations of the present worth of the future income which it is estimated may be obtained by their use; thus the price of pig-iron represents the discounted expected value, some months hence, of automobiles and other articles made from iron; the price of wheat is based upon the anticipated price of the bread which will eventually be made from the wheat; and the price of lumber represents a fraction of the present worth of anticipated house rents. Prices of goods sold for immediate consumption reflect conditions of demand and supply actually existing. Prices of indirect good represent expected conditions of demand and supply which may or may not come to pass. Prices of consumers' goods, being based upon actual conditions which

change but slowly, are relatively stable. On the other hand, prices at wholesale vary much more radically—the reason being that it is much easier to change guesses about the future than to change actual conditions of supply and demand.

Prices of stocks are removed from reality one step further than are the prices of commodities sold at wholesale, for they show present guesses concerning the value of goods in the future which it is expected will be sold at wholesale and retail—(mostly at wholesale). Depending as they do upon two layers of guesses, it is not surprising that stock prices fluctuate even more violently than do wholesale prices. All guesses concerning the future are likely to be far from the truth. Estimates made for a long period in advance tend, however, as time passes, to be steadily revised in the direction of the truth.

Since prices of direct goods constitute the ultimate goal toward which all other prices tend, it is not surprising to find that, though the prices of stocks and other indirect goods often diverge sharply from the relatively smooth trend representing the prices of direct or consumers' goods, they are almost certain to return eventually to the course of the prices of direct goods. The retail price index may, then, be thought of as the norm about which all other prices tend to fluctuate. Under such circumstances, there are strong arguments for adopting a comprehensive index of the prices of consumers' goods for use both as a standard of deferred payments and as the index to be used as a guide in stabilizing the dollar.

A general deflating index not obtainable.—As stated earlier in this chapter, one of the four reasons why economists are anxious to secure an index number measuring changes in the general price level is to have at hand a series of divisors for use in deflating values expressed in terms of current dollars. Presumably, it would be somewhat better to use as a deflator an inclusive index like that representing

the general purchasing power of money than it is to follow the practice so often adopted of utilizing for this purpose an index of wholesale prices. At best, however, the use of any price index for numerous purposes is likely to give results that are quite unsatisfactory. The object of the process of deflation is to arrive at quotients representing not values but physical quantities. To accomplish this end with any degree of precision, it is essential that the price-index number used should, in every case, be exactly adapted to the problem. If, for example, one has at hand statistics of the total value of exports from the United States and wishes to estimate therefrom the changes that have occurred in the physical quantities of goods exported, it is wholly unlikely that satisfactory results can be obtained through division by a price-index having weights which do not change during the entire period. The composition of our export trade varies greatly from month to month, and from year to year. To prove even moderately satisfactory as a divisor for reducing export totals in dollars to an index representing physical quantities of exports, a price index must take into account the variations in the type of exports which have actually occurred. It appears, then, that, for this specific purpose, the index number must be computed by the aid of variable weights, using a formula similar to that employed in calculating "P" in the Equation of Exchange. Such an index number would, however, in all probability, be entirely unsatisfactory for use in deflating total values of another kind.

Similarly, an index number suitable for reducing money wages to terms of constant purchasing power is satisfactory exactly to the extent to which the commodities covered by the price quotations used represent the average budget of the class of workers whose wages it is intended to "deflate." The best index for this purpose would probably give absurd results when used in an attempt to reduce some other class of money values to dollars of constant purchasing power.

It is, then, clearly futile to seek for an index series which will prove satisfactory, for general use as a "deflator."

Summary

1. The assumption that the only important use of index numbers is to measure changes in the purchasing power of money is decidedly erroneous. However, although it is probably impossible to measure with any high degree of accuracy, changes in the general purchasing power of money, it is, nevertheless, highly desirable to measure the changes as accurately as possible.

2. The best index number for this purpose is one which

(a) Includes a wide list of commodities in different fields.
(b) Uses a dollar's worth of every commodity in the base period as one unit.
(c) Varies the quantities from period to period as production and consumption change.
(d) Consists of a combination of an index of trend with an index of short-time fluctuations.

3. Prices of common stocks should be included in the short-time index, but not in the trend index.

4. The index of the general purchasing power of money constitutes the best medium for measuring the effect upon prices of changes in the forces affecting the supply of and the demand for circulating medium.

5. The movement for stable money has intensified the demand for an index especially designed for use as a standard for measuring deferred payments. Although many persons believe that such an index should be based upon prices of commodities of all sorts, it appears to be most logical to use in its calculation prices of direct goods only.

QUESTIONS AND PROBLEMS

1. In the phrase "the general purchasing power of money," what is meant by the term

(a) Money?
(b) General purchasing power?

2. Should the following items be listed under "trade" in the equation of exchange?

(a) Loaning $10 to a friend.
(b) Loaning $5,000 on a mortgage.
(c) Buying a bond.

Are these transactions essentially different in nature?

3. What date-differences interfere with the accurate measurement of the quantities entering into the equation of exchange?

4. Why are the following not suitable for inclusion among the commodities used in making up an index of the *trend* of the general price-level?

(a) House rents.
(b) Corporate stocks.
(c) Bonds.
(d) Automobiles.
(e) Real estate.

5. Is the meaning of a series of relatives derived from complete data easier to understand when the weights used are fixed or when they are variable?

6. Compare the merits of fixed weights and variable weights when employed in the computation of an index series designed for use as a standard of deferred payments.

7. How would you go about it to combine a cyclical price index with a trend derived from another price index?

8. As a standard of deferred payments, which is preferable, a general price index, or an index of the prices of direct goods? Why?

9. Ought the commodity value of a debt to increase in proportion to the growing productivity of industry?

10. Would it be desirable to have a standard of deferred payments based upon wages per hour—the object being to make a given debt always represent the same number of hours of labor of average quality?

11. Why is it not possible to devise a price index which can be used as a general deflator?

BIBLIOGRAPHICAL NOTE

For discussions of the problem of constructing index numbers for measuring the general purchasing power of money, see:

F. Y. Edgeworth, *Journal of the Royal Statistical Society,* December, 1883, pp. 714–718; Reports of the Committee of the British Association for the Advancement of Science on *The Best Methods of Ascertaining and Measuring Variations in the Value of the Monetary Standard,* pp. 188–219 (1888); and *Journal of the Royal Statistical Society,* July, 1925, pp. 557–575.

Correa M. Walsh, *The Measurement of General Exchange Value* (1901), whole book, and *The Quarterly Journal of Economics,* May, 1924, pp. 500–519.

Irving Fisher, *The Purchasing Power of Money* (1911), whole book —especially pp. 198–225.

G. H. Knibbs, Australian Commonwealth Bureau of Census and Statistics, Labor and Industrial Branch, *Report No. 9,* July, 1919.

A. C. Pigou, *The Economics of Welfare* (1920), pp. 69–90.

George R. Davies, *Journal of the American Statistical Association,* June, 1924, pp. 180–188.

Hudson B. Hastings, *Journal of the American Statistical Association,* Papers and Proceedings, March, 1928, pp. 142–145.

CHAPTER IX

CONCLUSION

The first purpose of this book.—One of the prime purposes of this book, as stated in the Introduction, is to get rid of the mathematical haze surrounding the subject of index numbers, and to show in extremely simple form what index numbers mean, how they are to be interpreted, how they differ from one another, when each of the most common types is to be used, and how each such type is to be constructed.. It is hoped that material progress has been made in each of these directions.

The second purpose of this book.—The second main reason for writing this book has been to explode a number of fallacies which have gained wide acceptance. It is hoped, therefore, that no reader of this volume will, in the future, make any of the following fallacious assumptions:

1. That the only worth-while reason for constructing index numbers is to measure the general purchasing power of money.
2. That there is an index number which is best for all purposes.
3. That the merit of an index series can be judged by any of the following criteria:
   (a) How closely it follows the central tendency of a group of index numbers—an assumption no more logical than that of taking the central location of a charge of bird shot fired into a target as being the bull's-eye at which the shot was fired.
   (b) The circular test.
   (c) The factor-reversal test.
   (d) The commodity-reversal test.
   (e) The time-reversal test.

4. That the formula applicable to the totality is necessarily the best formula for use in constructing an index number from a small sample.

Finally, it is hoped that the fact has been made absolutely clear that the problem of index numbers is merely a specific instance of the problem of sampling, and that, therefore, the merit of an index number depends wholly upon how closely the average of the sample corresponds to the average of the totality.

The third purpose of this book.—The third fundamental reason for inflicting another book upon the public has been to present the results of experiments in this field made with concrete data. It is hoped that most students will find useful the figures showing the percentages of error found in various index numbers when these indices are used in an attempt to answer one or another of three specific questions. Perhaps, also, the studies of the peculiarities of index numbers computed from chain relatives, and the investigations into methods of reducing the size of errors in index numbers, may be helpful to workers in this field.

While the author likes to believe that he has made progress in each of the three directions mentioned, the final verdict necessarily must be given by the reader. Obviously, the subject of index numbers is one having so many ramifications that it is impossible, in a single brief volume, to do more than to touch upon some of the main points. If the book has succeeded in setting on the correct path the student who expects to do serious work in this field, it has achieved all of its aims.

INDEX

Absolute comparisons, 5
Absolute numbers,
  conversion to relatives, 5
Accuracy of sample items, 60–61
Aggregates, method of, viii, 18–20, 24
  applied to stock prices, 116–124, 146–161
  dominating effect of large items on, 31–32
  equals av. of relatives, 25, 31–34, 85–86
  misuse of, 70–71
  use with changing sample, 86–88
  weighted, 68, 71, 84, 85
American Economic Review, 4, 89, 91, 105
Answers to questions,
  asked, 160–161, 163, 183–184
  different relatives answer different, 19–31, 25–31, 49–53
  exact possible with complete data, 40–41
  formulation of, 7–8
  necessity of making them precise, 7–8, 19–31, 49–53, 146, 197–198
  often vague, 7–8
Arithmetic average
  compared with other averages, 108–140, 130–132
  of price relatives, 162–170, 168–180
  use as base, 11–18
  weighted compared with unweighted, 131–142, 163–164, 170–173, 176–187
Ashby, Forrest B., 142
Averages of relatives, viii, 18–19, 21–22, 25, 32–35, 69–76, 81–84
Averages of relatives affected by weights chosen, 127–129
  of commodity prices, 161–187
  weighted and unweighted compared, 130–142, 163–164, 172–173, 176–187

Bagwell, Lucile, xi
Base,
  choice of, 8–9, 11–18
  definition of, 8–9
  effect of different, 11–18
  effect of shifting, 34–35
  for general price index, 196–198
  price and weight bases different, 35–36
  relationship to per cent of error, 150–151, 165–166, 185
  shift in, changes weights, 23–25
Bean, L. H., 44, 58, 142
Behavior of Prices, 151
Belcher, Donald R., 4
Bias of index numbers
  actual, 87–94
  supposed existence of, ix, 126–128, 140
Bogart, Ernest L., xi
Bond prices
  relation to general price index, 206–209
Bowley, Arthur L., 4, 44, 142, 188
Bureau of Labor Statistics, xi, 4, 17, 58, 130, 161
  Bulletin 284, xi, 4, 44, 58, 105, 142
  Bulletin 390, 17

Carriages and motor vehicles, 11–14
Central tendency, ix

Chaddock, Robert E., 43–44, 77
Chained relatives, 16–18, 78–95, 97–99, 203–204
Charts and Graphs, 43
Choice of index number, 2
Circular test, ix, 55, 219
City and hamlet, 5–6
Civil Service, 73–75
Cloth and cotton, 7–9
Commodities
  prices of, 100–105
  substitution in price comparisons, 38–40, 79–93, 203–204
Commodity-prices,
  averages of relatives, 161–187
  indices of, 129–142
  samples selected, 129–131
Commodity-reversal test, ix, 55, 219
Comparisons
  commodities must be identical, 36–38
  distant ones difficult, 37–38
  on different bases, 18–25
  relative vs. absolute, 5
  substitution of commodities, 38–40, 79–93, 203–204
  using complete data, 5–44
Complete data, 55
  comparisons using, 5–44
  vs. sample data, 59
Consumers' goods,
  price of as standard of payments, 213–216
Copeland, Morris A., 192
Corn and cloth, 7–9
Corn and cotton, 27–29
Crop Reporters, 60, 65–73
Crump, Norman, 58, 105, 142
Cyclical index
  compared with trend index, 203–209

Davies, George R., 44, 218
Deflator, general
  non-existent, 214–216
Department of Commerce, 9
Dessel, Rose, xi
Direct goods
  price of as standard of payments, 213–216
Direct vs. indirect services, 206–209
Dobson, Harriet, xi
Dublin, Betty, xi

Economic Journal, 4
Economics of Welfare, 142, 218
Edgeworth, F. Y., 4, 142, 188, 190, 218
Encyclopaedia of Index Numbers, xi
Equation of exchange, 192–202, 215
Erratic items,
  increase errors, 176–177, 182–187
Error, percentage of
  in index numbers of
    commodity prices, 163–187
    stock prices, 145–164

Factor-reversal test, ix, 55, 219
Factory wages, 29–30
Facts and Figures of the Automobile Industry, 10
Federal Reserve Bulletin, 58
Field, James A., 43
Fisher, Irving, xi, 3, 43–44, 58, 94, 105, 113, 126–128, 132, 140–141, 189–190, 192, 218
Flinn, Harold M., 4
Fork, three-tined, 113–116
Formula for
  general price index, 201–208
  index numbers, 143–146
Formulation of questions, 7–8
  necessity of precision in, 19–31, 49–53
Fountain, H., 142

Geometric average, 108–118, 124–125
  of relatives of commodity prices, 131–132, 163–165, 170

Hamlet and city, 5–6
Handbook of Mathematical Statistics, 77
Harmonic average, 108–116
Hastings, Hudson B., 218

High School grades
comparison of, 19–26
House rent, 204–205, 209

Ideal formula for index numbers, ix, 2–3
universal non-existent, 53–55
Importance of commodities
effect on weighting, 173–177
Income
per capita, 10
Income in the Various States, 10
Index number
choice of, 2
defined, 6
splicing of, 71–76, 80–82
Index numbers
based upon sample data, 46–47, 49
defined, 47–49
Encyclopaedia of, xi
erroneous views concerning, viii–x
established facts concerning, 144
formula applicable to totality, 144
historical development of, 4
nature of, 45–58
no universal formula possible, 53–55
percentages of error in, 143–187
problem of, ix–x, 2, 49
questions concerning, vii–viii
supposed bias of, 89–95, 113–116, 126–128
tests for, 53–58, 111–114
tests for accuracy, 146–181
variety of relatives, 46–48
Index numbers of
commodity prices at wholesale, 126–142, 161–187
general price level, 188–217
stock prices, 116–126, 145–161

Jacobson, Bertha, xi
Jerome, Harry, 77
Journal of American Statistical Association, 4, 44, 58, 141–142, 192, 218
Journal of Political Economy, 43, 105, 142, 199
Journal of Royal Statistical Society, 4, 58, 105, 142, 188, 218

Karsten, Karl G., 43
Knibbs, George H., 4, 44, 57, 141–142, 218
Kresge stock, 99

Labor, price of, 205–206, 208
as standard of deferred payments, 209–213
Les Nombres Indices de la Variation des Prix, 4
Leven, Maurice, xi, 183, 210
Levin, Benjamin S., xi
Link relatives, 78–105
splicing of averages of, 86–94, 203–204

Macaulay, Frederick R., 89, 91, 105, 141
Making of Index Numbers, 44, 58, 94, 105, 113, 126–127, 141, 189–190
Mean
(*see* average)
Measurement of General Exchange Value, 142, 218
Median, 107–108, 124–125, 148–150
of relatives of commodity prices, 131–132, 163–165, 170
Merit of an index series, ix–x, 55–58
Method of aggregates, viii, 18–20, 24
applied to stock prices, 116–124, 146–161
dominating effect of large items on, 31–32
equals average of relatives, 25, 31–34, 85–86
misuse of, 70–71
use with changing sample, 86–88
weighted, 68–71, 84–85

Mills, Frederick C., xi, 77, 151
Mitchell, Wesley C., xi, 3–4, 43–44, 58, 105, 142, 199
Money
  purchasing power of, viii, 188–217
  uses of, 191–192
Monthly Labor Review, 58
Moody's *Rating-Book*, 95
Motor vehicle ownership, 10
Motor vehicles
  and carriages, 11–14

National Automobile Chamber of Commerce, 10
National Bureau of Economic Research, xi, 10, 89, 151

Olivier, Maurice, 3–4

Papers Relating to Political Economy, 190
Payne, Charles K., xi
Peabody, Walter R., xi
Percentages of error in
  commodity price indices, 163–187
  stock price indices, 145–164
Physical production
  measurement of, 27–29
Physical volume of trade, 194–200
Piece rates vs. time rates, 205–206, 208
Pierson's time-reversal test, ix, 55, 111–114, 219
Pigou, A. C., 142, 218
Price data
  accuracy of, 61
  averages of, 61–62
Price indices of
  commodities at wholesale, 126–142, 161–187
  general price level, 188–217
  stock prices, 116–126, 145–161
Price level
  measurement of variations in, viii, 188–217
Price relatives
  choice of base, 13–18
  commodity, 161–187
Price relatives
  commodity substitution in, 38–40
  must compare same things, 36–38
  splicing of, 39–40, 71–76, 80–82, 86–94
Prices of
  commodities, 100–103
  labor, 205–206, 209–213
  stocks, 95–99
Principles and Methods of Statistics, 43–44, 77
Problem of Estimation, 189
Proportional vs. random sampling, 64–67, 181–182, 187
Purchasing power of money, viii, 188–217
Purchasing Power of Money, 44, 192, 218
Purposes of this book, vii–x, 1–3, 219–220

Quarterly Journal of Economics, 4, 44, 218
Quarterly Publications of American Statistical Assn., 43
Questions answered by this book, vii–x
Questions
  answers to, 160–161, 163, 183–184
  asked, 146, 162, 168
  necessity of making precise, 7–8, 19–31, 49–53, 146, 197–198
  vague, 7–8

Random sampling
  vs. proportional, 64–67, 181–182, 187
Ratio charts, 18
Relative Comparisons, 5
  commodities must be identical, 36–38
  distant comparisons difficult, 37–38
  on different bases, 18–25
  substitution of commodities, 38–40
  usefulness not confined to time series, 10

Relatives
advantages over absolute numbers, 7
averages of, viii, 18–19, 21–23, 69–76, 81–86
equal to weighted aggregate, 25, 32–35
based upon complete data, 46–47
chain, 16–18
defined, 6
different for different questions, 19–31
fixed base, 97–99
link, 78–81, 97–99
of aggregates, 116–124
of commodity prices, 161–187
of stock prices, 111–127, 145–160
splicing of, 39–40, 71–76, 80–82, 86–94
weakness of fixed base, 16
weighted average equal to aggregate, 86
Rent of houses, 204–205, 209
Rietz, H. L., 77

Sample
changing, 74–76, 80–81
introducing new items into, 71–73
number of items needed, 62–64
removing erratic items from, 182–187
size of as related to size of error, 177–181, 186
weighting of items in, 67–73
Sample data
necessarily used for general price index, 200
Samples
of commodity prices, 100–101, 128–130
of stock prices, 96–97
size as related to per cent of error, 151–157, 166–168
Sampling
as related to index numbers, 59–76
effects of poor, ix
Sampling
index numbers based upon, ix, 2–3
random vs. proportional, 64–67, 181–187
Scope of book, vii–x, 1–3
Services and general price index, 204–209
Simple arithmetic average
compared with weighted, 131–142, 163–164, 170–173, 176–187
of price relatives, 162–170
Snyder, Carl, 4
Splicing of relatives or index numbers, 39–40, 71–76, 80–82, 86–94
Standardization of commodities
lack of hinders index making, 199–200
Standard of deferred payments, 209–214
Stine, Oscar C., 44, 58, 142
Stock,
share not a fixed unit, 199
Stock-price index numbers, 116–126
percentages of error in, 145–164
Stock prices, 95–99
aggregate values, 31–33
relation to general price index, 207–209
Substitution of commodities, 38–40, 79–93, 203–204
Sugar, relative price of, 17
Survey of Current Business, 9

Tests for index numbers, ix–x, 53–58, 111–114, 219
Time-reversal test, ix, 55, 111–114, 219
Thomas, Woodlief, 142
Totality, ix, 144–146, 162, 169
Trade,
physical volume of, 194–200
Trend index
compared with cyclical index, 203–209
Types and averages, 108–140

Unfilled orders of U. S. Steel Corp., 9
Units for measuring
physical production, 28–29
trade, 195–200
Universe of items, ix, 145–146
U. S. Bureau of Labor Statistics, xi, 4, 17, 58, 130, 161
Bulletin 284, xi, 4, 44, 58, 105, 142
Bulletin 390, 17
U. S. Department of Commerce, 9
U. S. Steel Corporation
earnings and unfilled orders, 9

Wages, 205–206
as standard of deferred payments, 209–213
factory, 29–30
Walsh, Correa M., 4, 142, 189, 218
Weighted arithmetic average of relatives
compared with unweighted average, 131–142, 163–164, 172–173, 176–187
equal to method of aggregates, 85–86
of commodity prices, 168–180
Weighting, 72–73
change in base equivalent to, 23–25, 115–116
effect of changing weights, 21–25, 127–129
individual commodity vs. group, 173–176
necessary in computing general price index, 200–202
results when numbers are converted to relatives, 115–116
why it sometimes increases error of index, 171–173
Wholesale Prices
Bulletin 390, U. S. Bur. of Labor Stat., p. 17 (*see also* "Commodities" and "Commodity prices")
Young, Allyn A., 44
Yule, G. Udny, 4

Bookstore 1-18-35

I S U LIBRARY
3 2792 008 242 574

WITHDRAWN
IOWA STATE UNIVERSITY
LIBRARY

Iowa State University Library
Ames, Iowa 50010